

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US05/010913

International filing date: 31 March 2005 (31.03.2005)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/559,011
Filing date: 01 April 2004 (01.04.2004)

Date of receipt at the International Bureau: 12 August 2005 (12.08.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

1352380

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

August 02, 2005

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: 60/559,011

FILING DATE: April 01, 2004

RELATED PCT APPLICATION NUMBER: PCT/US05/10913



Certified by

Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office

01919 U.S. PTO
040104

PTO/SB/16 (08-03)
Approved for use through 7/31/2006. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No. EL 961008135 US

INVENTOR(S)					
Given Name (first and middle (if any))		Family Name or Surname		Residence (City and either State or Foreign Country)	
Steven Andreas Stefan M.		MAH BRAUN KAMMERER		San Diego, California San Diego, California San Diego, California	
Additional inventors are being named on the <u>1</u> separately numbered sheets attached hereto					
TITLE OF THE INVENTION (500 characters max)					
METHODS FOR IDENTIFYING RISK OF OSTEOARTHRITIS AND TREATMENTS THEREOF					
Direct all correspondence to: CORRESPONDENCE ADDRESS					
<input checked="" type="checkbox"/> Customer Number:		25225			
OR					
<input type="checkbox"/> Firm or Individual Name					
Address					
City		State		Zip	
Country		Telephone		Fax	
ENCLOSED APPLICATION PARTS (check all that apply)					
<input checked="" type="checkbox"/>	Specification Number of Pages	214	<input type="checkbox"/>	CD(s), Number	
<input checked="" type="checkbox"/>	Drawing(s) Number of Sheets	2	<input checked="" type="checkbox"/>	Other	Return Receipt Postcard
<input checked="" type="checkbox"/>	Application Data Sheet. See 37 CFR 1.76 (4 pages) (specify):				
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT					
<input checked="" type="checkbox"/>	Applicant claims small entity status. See 37 CFR 1.27.				FILING FEE AMOUNT (\$) 80.00
<input type="checkbox"/>	A check or money order is enclosed to cover the filing fees.				
<input checked="" type="checkbox"/>	The Director is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number: 03-1952				
<input type="checkbox"/>	Payment by credit card. Form PTO-2038 is attached.				
The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.					
<input checked="" type="checkbox"/>	No	<input type="checkbox"/> Yes, the name of the U.S. Government agency and the Government contract number are:			

[Page 1 of 2]

Respectfully submitted,

Date April 1, 2004

SIGNATURE
TYPED OR
PRINTED NAME
TELEPHONE

Bruce D. Grant
(858) 720-7962

REGISTRATION NO.
(if appropriate) 47,608
Docket Number: 524593008600

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

I hereby certify that this correspondence is being deposited with the U.S. Postal Service as Express Mail, Airbill No. EL 961008135 US, in an envelope addressed to: Mail Stop, Provisional Patent Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on the date shown below.

Dated: 4/1/04 Signature: (Deborah Wykes)

SD-189414

PROVISIONAL APPLICATION COVER SHEET
Additional Page

PTO/SB/16 (08-03)

Approved for use through 07/31/06. OMB 0651-0032

U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Docket Number 524593008600

INVENTOR(S)/APPLICANT(S)		
Given Name (first and middle [if any])	Family or Surname	Residence (City and either State or Foreign Country)
Matthew Roberts	NELSON	San Marcos, California
Rikard Henry	RENELAND	San Diego, California
Maria	LANGDOWN	San Diego, California

[Page 2 of 2]

METHODS FOR IDENTIFYING RISK OF OSTEOARTHRITIS AND TREATMENTS THEREOF

Field of the Invention

[0001] The invention relates to genetic methods for identifying risk of osteoarthritis and treatments that specifically target such diseases.

Background

[0002] Osteoarthritis (OA) is a chronic disease usually affecting weight-bearing synovial joints. There are approximately 20 million Americans affected by OA and it is the leading cause of disability in the United States. In addition to extensive human suffering, OA also accounts for nearly all knee replacements and more than half of all hip replacements in the United States. Despite its prevalence, OA is poorly understood and there are few treatments available besides anti-inflammatory drugs and joint replacement.

[0003] Most commonly affecting middle-aged and older people, OA can range from very mild to very severe. It affects hands and weight-bearing joints such as knees, hips, feet and the back. Knee OA can be as disabling as any cardiovascular disease except stroke.

[0004] OA is characterized by the breakdown of cartilage in joints. Cartilage in joints cushions the ends of bones, and cartilage breakdown causes bones to rub against each other, causing pain and loss of movement. Type II collagen is the main component of cartilage, comprising 15-25% of the wet weight, approximately half the dry weight, and representing 90-95% of the total collagen content in the tissue. It forms fibrils that endow cartilage with tensile strength (Mayne, R. Arthritis Rheum. 32:241-246 (1989)).

Summary

[0005] It has been discovered that certain polymorphic variations in human genomic DNA are associated with osteoarthritis. In particular, polymorphic variants in loci containing *KIAA0296*, *chrom 4* and *SNW1* regions and other regions in Table A of human genomic DNA have been associated with risk of osteoarthritis.

[0006] Thus, featured herein are methods for identifying a subject at risk of osteoarthritis and/or a risk of osteoarthritis in a subject, which comprise detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in or around the loci described herein in a human nucleic acid sample. In an embodiment, two or more polymorphic variations are detected in two or more regions of which one is the *KIAA0296*, *chrom 4* or *SNW1* region or other region in Table A. In certain

embodiments, 3 or more, or 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 or more polymorphic variants are detected.

[0007] Also featured are nucleic acids that include one or more polymorphic variations associated with occurrence of osteoarthritis, as well as polypeptides encoded by these nucleic acids. In addition, provided are methods for identifying candidate therapeutic molecules for treating osteoarthritis, as well as methods for treating osteoarthritis in a subject by identifying a subject at risk of osteoarthritis and treating the subject with a suitable prophylactic, treatment or therapeutic molecule.

[0008] Also provided are compositions comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and/or a *KIAA0296*, *chrom 4* or *SNW1* nucleic acid or other nucleic acid referenced in Table A, with a RNAi, siRNA, antisense DNA or RNA, or ribozyme nucleic acid designed from a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A. In an embodiment, the RNAi, siRNA, antisense DNA or RNA, or ribozyme nucleic acid is designed from a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A that includes one or more polymorphic variations associated with osteoarthritis, and in some instances, specifically interacts with such a nucleotide sequence. Further, provided are arrays of nucleic acids bound to a solid surface, in which one or more nucleic acid molecules of the array have a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A, or a fragment or substantially identical nucleic acid thereof, or a complementary nucleic acid of the foregoing. Featured also are compositions comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and/or a *KIAA0296*, *chrom 4* or *SNW1* polypeptide or other polypeptide referenced in Table A, with an antibody that specifically binds to the polypeptide. In an embodiment, the antibody specifically binds to an epitope in the polypeptide that includes a non-synonymous amino acid modification associated with osteoarthritis (e.g., results in an amino acid substitution in the encoded polypeptide associated with osteoarthritis). In certain embodiments, the antibody selectively binds to an epitope in the *KIAA0296*, *chrom 4* or *SNW1* polypeptide, or other polypeptide referenced in Table A, having an amino acid associated with osteoarthritis. Thus, featured is an antibody that binds an epitope having an amino acid encoded by rs734784, rs1042164, rs749670, rs955592 and/or rs1040461, such as a valine or isoleucine encoded by rs734784 (e.g., a valine at position 489 in a *KCNS1* polypeptide), a valine or alanine encoded by rs1042164 (e.g., a valine at position 133 in a *ETR101* polypeptide), a glutamate or glycine encoded by rs749670 (e.g., a glutamate at position 327 in a *KIAA0296* polypeptide), a threonine or isoleucine encoded by rs955592 (e.g., a threonine at position 70 in a *FLJ21977* polypeptide) or a glycine or serine encoded by rs1040461 (e.g., a glycine at position 207 in a *RAB23* polypeptide) at the corresponding position in the polypeptide.

Brief Description of the Drawings

[0009] Figures 1A-1C show proximal SNPs in *KIAA0296*, *chrom 4* and *SNW1* regions of genomic DNA, respectively. The position of each SNP in the chromosome is shown on the x-axis and the y-axis provides the negative logarithm of the p-value comparing the estimated allele to that of the control group. Also shown in the figures are exons and introns of the regions in the approximate chromosomal positions.

Detailed Description

[0010] It has been discovered that a polymorphic variant in a locus containing a *KIAA0296*, *chrom 4* or *SNW1* region is associated with occurrence of osteoarthritis in subjects. Thus, detecting genetic determinants associated with an increased risk of osteoarthritis occurrence can lead to early identification of a predisposition to osteoarthritis and early prescription of preventative measures. Also, associating a *KIAA0296*, *chrom 4* or *SNW1* polymorphic variant and other variants referenced in Table A with osteoarthritis has provided new targets for screening molecules useful in treatments of osteoarthritis.

Osteoarthritis and Sample Selection

[0011] Osteoarthritis (OA), or degenerative joint disease, is one of the oldest and most common types of arthritis. It is characterized by the breakdown of the joint's cartilage. Cartilage is the part of the joint that cushions the ends of bones, and its breakdown causes bones to rub against each other, causing pain and loss of movement. Type II collagen is the main component of cartilage, comprising 15-25% of the wet weight, approximately half the dry weight, and representing 90-95% of the total collagen content in the tissue. It forms fibrils that endow cartilage with tensile strength (Mayne, R. Arthritis Rheum. 32:241-246 (1989)).

[0012] Most commonly affecting middle-aged and older people, OA can range from very mild to very severe. It affects hands and weight-bearing joints such as knees, hips, feet and the back. Knee OA can be as disabling as any cardiovascular disease except stroke.

[0013] Osteoarthritis affects an estimated 20.7 million Americans, mostly after age 45, with women more commonly affected than men. Physicians make a diagnosis of OA based on a physical exam and history of symptoms. X-rays are used to confirm diagnosis. Most people over 60 reflect the disease on X-ray, and about one-third have actual symptoms.

[0014] There are many factors that can cause OA. Obesity may lead to osteoarthritis of the knees. In addition, people with joint injuries due to sports, work-related activity or accidents may be at increased risk of developing OA.

[0015] Genetics has a role in the development of OA. Some people may be born with defective cartilage or with slight defects in the way that joints fit together. As a person ages, these defects may

cause early cartilage breakdown in the joint or the inability to repair damaged or deteriorated cartilage in the joint.

[0016] Inclusion or exclusion of samples for an osteoarthritis pool may be based upon the following criteria: ethnicity (e.g., samples derived from an individual characterized as Caucasian); parental ethnicity (e.g., samples derived from an individual of British paternal and maternal descent); relevant phenotype information for the individual (e.g., case samples derived from individuals diagnosed with specific knee osteoarthritis (OA) and were recruited from an OA knee replacement clinic). Control samples may be selected based on relevant phenotype information for the individual (e.g., derived from individuals free of OA at several sites (knee, hand, hip etc)); and no family history of OA and/or rheumatoid arthritis. Additional phenotype information collected for both cases and controls may include age of the individual, gender, family history of OA, diagnosis with osteoarthritis (joint location of OA, date of primary diagnosis, age of individual as of primary diagnosis), knee history (current symptoms, any major knee injury, meniscectomy, knee replacement surgery, age of surgery), HRT history, osteoporosis diagnosis.

[0017] Based in part upon selection criteria set forth above, individuals having osteoarthritis can be selected for genetic studies. Also, individuals having no history of osteoarthritis often are selected for genetic studies, as described hereafter.

Polymorphic Variants Associated with Osteoarthritis

[0018] A genetic analysis provided herein linked osteoarthritis with polymorphic variant nucleic acid sequences in the human genome. As used herein, the term “polymorphic site” refers to a region in a nucleic acid at which two or more alternative nucleotide sequences are observed in a significant number of nucleic acid samples from a population of individuals. A polymorphic site may be a nucleotide sequence of two or more nucleotides, an inserted nucleotide or nucleotide sequence, a deleted nucleotide or nucleotide sequence, or a microsatellite, for example. A polymorphic site that is two or more nucleotides in length may be 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or more, 20 or more, 30 or more, 50 or more, 75 or more, 100 or more, 500 or more, or about 1000 nucleotides in length, where all or some of the nucleotide sequences differ within the region. A polymorphic site is often one nucleotide in length, which is referred to herein as a “single nucleotide polymorphism” or a “SNP.”

[0019] Where there are two, three, or four alternative nucleotide sequences at a polymorphic site, each nucleotide sequence is referred to as a “polymorphic variant” or “nucleic acid variant.” Where two polymorphic variants exist, for example, the polymorphic variant represented in a minority of samples from a population is sometimes referred to as a “minor allele” and the polymorphic variant that is more prevalently represented is sometimes referred to as a “major allele.” Many organisms possess a copy of each chromosome (e.g., humans), and those individuals who possess two major alleles or two minor

alleles are often referred to as being “homozygous” with respect to the polymorphism, and those individuals who possess one major allele and one minor allele are normally referred to as being “heterozygous” with respect to the polymorphism. Individuals who are homozygous with respect to one allele are sometimes predisposed to a different phenotype as compared to individuals who are heterozygous or homozygous with respect to another allele.

[0020] In genetic analysis that associate polymorphic variants with osteoarthritis, samples from individuals having osteoarthritis and individuals not having osteoarthritis often are allelotyped and/or genotyped. The term “allelotype” as used herein refers to a process for determining the allele frequency for a polymorphic variant in pooled DNA samples from cases and controls. By pooling DNA from each group, an allele frequency for each SNP in each group is calculated. These allele frequencies are then compared to one another. The term “genotyped” as used herein refers to a process for determining a genotype of one or more individuals, where a “genotype” is a representation of one or more polymorphic variants in a population.

[0021] A genotype or polymorphic variant may be expressed in terms of a “haplotype,” which as used herein refers to two or more polymorphic variants occurring within genomic DNA in a group of individuals within a population. For example, two SNPs may exist within a gene where each SNP position includes a cytosine variation and an adenine variation. Certain individuals in a population may carry one allele (heterozygous) or two alleles (homozygous) having the gene with a cytosine at each SNP position. As the two cytosines corresponding to each SNP in the gene travel together on one or both alleles in these individuals, the individuals can be characterized as having a cytosine/cytosine haplotype with respect to the two SNPs in the gene.

[0022] As used herein, the term “phenotype” refers to a trait which can be compared between individuals, such as presence or absence of a condition, a visually observable difference in appearance between individuals, metabolic variations, physiological variations, variations in the function of biological molecules, and the like. An example of a phenotype is occurrence of osteoarthritis.

[0023] Researchers sometimes report a polymorphic variant in a database without determining whether the variant is represented in a significant fraction of a population. Because a subset of these reported polymorphic variants are not represented in a statistically significant portion of the population, some of them are sequencing errors and/or not biologically relevant. Thus, it is often not known whether a reported polymorphic variant is statistically significant or biologically relevant until the presence of the variant is detected in a population of individuals and the frequency of the variant is determined. Methods for detecting a polymorphic variant in a population are described herein, specifically in Example 2. A polymorphic variant is statistically significant and often biologically relevant if it is represented in 5% or more of a population, sometimes 10% or more, 15% or more, or 20% or more of a population, and often 25% or more, 30% or more, 35% or more, 40% or more, 45% or more, or 50% or more of a population.

[0024] A polymorphic variant may be detected on either or both strands of a double-stranded nucleic acid. Also, a polymorphic variant may be located within an intron or exon of a gene or within a portion of a regulatory region such as a promoter, a 5' untranslated region (UTR), a 3' UTR, and in DNA (*e.g.*, genomic DNA (gDNA) and complementary DNA (cDNA)), RNA (*e.g.*, mRNA, tRNA, and rRNA), or a polypeptide. Polymorphic variations may or may not result in detectable differences in gene expression, polypeptide structure, or polypeptide function.

[0025] It was determined that polymorphic variations associated with an increased risk of osteoarthritis existed in SEQ ID NO: 1-3 or a nucleotide sequence referenced in Table A. In certain embodiments, polymorphic variants at positions rs552, rs12904, rs2282146, rs734784, rs1042164, rs749670, rs955592, rs1143016, rs755248, rs1055055, rs835409, rs927663, rs8162, rs831038, rs33079, rs1710880, rs4472029, rs799570, rs1282730, rs1518875, rs1568694, rs905042, rs1957723, rs794018, rs707723, rs893861, rs1914903, rs2062232, rs26609, rs1370987, rs1012414, rs435903, rs1248, rs703508, rs226465, rs241448, rs763155, rs1040461, rs462832, rs804194, rs1022646, rs1569112, rs805623, rs1019850, rs1599931, AA, rs279941, rs1062230, rs1859911, rs1477261, rs1191119, rs657780, rs1393890, rs1478714, rs868213, rs690115, rs1465501, rs899173, rs10477, rs926393, rs465271, rs13847 and/or rs738658 in the human genome were associated with an increased risk of osteoarthritis, and in specific embodiments, the corresponding allele in the right-most column in Table A for each position is associated with an increased risk of osteoarthritis. In other embodiments polymorphic variants at positions rs734784, rs1042164, rs749670, rs955592 and rs1040461 were associated with an increased risk of osteoarthritis, and in specific embodiments, a valine encoded by rs734784, a valine encoded by rs1042164, a glutamate encoded by rs749670, a threonine encoded by rs955592 and a glycine encoded by rs1040461 were associated with an increased risk of osteoarthritis.

[0026] Polymorphic variants in and around the *KIAA0296* locus were tested for association with osteoarthritis. These include polymorphic variants at positions in SEQ ID NO: 1 selected from the group consisting of 247, 1535, 2386, 6440, 9133, 9143, 9471, 13150, 13717, 14466, 15769, 16870, 18545, 18749, 19123, 20736, 21038, 21046, 21050, 21056, 21706, 23170, 25028, 27871, 28070, 31717, 32019, 32318, 33080, 33101, 34236, 34285, 34818, 35168, 37981, 38113, 38117, 38481, 38615, 38944, 39288, 41385, 42136, 42185, 42353, 42434, 44580, 44675, 45739, 46439, 47457, 47735, 50319, 50708, 51185, 53002, 53064, 53637, 55274, 55825, 55986, 56684, 57653, 57659, 57692, 57775, 61313, 61431, 61699, 62906, 63619, 64664, 68452, 69665, 69681, 70091, 74637, 74760, 76523, 78559, 79549, 79882, 81339, 81681, 81696, 83517, 85431, 86332, 87358, 87725, 89052, 90020, 90231, 90284, 90447, 90601, 90724, 92559, 95176, 95195 and 96822. Polymorphic variants at the following positions in SEQ ID NO: 1 in particular were associated with an increased risk of osteoarthritis: 13150, 21046, 23170, 25028, 44580, 62906, 64664 and 83517. In particular, the following polymorphic variants in SEQ ID NO: 1 were associated with risk of osteoarthritis: a guanine at position 13150, a thymine at position 21046, an

adenine at position 23170, an adenine at position 25028, a guanine at position 44580, a guanine at position 62906, a cytosine at position 64664 and a cytosine at position 83517. A polymorphic variant in a *KIAA0296* polypeptide encoded by rs749670 (e.g., a glutamate at position 327 in the polypeptide) also was associated with increased risk of osteoarthritis.

[0027] Polymorphic variants in and around the *chrom 4* locus were tested for association with osteoarthritis. These include polymorphic variants at positions in SEQ ID NO: 2 selected from the group consisting of 211, 7217, 7895, 13308, 14279, 17026, 18271, 20417, 21843, 22069, 22145, 22519, 22539, 23236, 23256, 23402, 23499, 23620, 23871, 24136, 25427, 25866, 26541, 26576, 26689, 26720, 27113, 27164, 27186, 28341, 29160, 29844, 30665, 30830, 31061, 31523, 32326, 32346, 32358, 34909, 34975, 35066, 35096, 35375, 36304, 36712, 36770, 37342, 37412, 37884, 38077, 38300, 38301, 41189, 44408, 44493, 44571, 44670, 45219, 45258, 47261, 48473, 48771, 55292, 56479, 56747, 60620, 60688, 61058, 61129, 61577, 61961, 63351, 63926, 65798, 66043, 66044, 66246, 66318, 66547, 71238, 71283, 71492, 72274, 73762, 74209, 75284, 77347, 77589, 78096, 78606, 78862, 79135, 79146, 79456, 79609, 80086, 80119, 80766, 81110, 81269, 81668, 82433, 82559, 83298, 83821, 84121, 84147, 84543, 84554, 84691, 84727, 85678, 86699, 86700, 86792, 86832, 87045, 87140, 87365, 88342, 88498, 88589, 95502, 96968, 97448, 97568 and 98724. Polymorphic variants at the following positions in SEQ ID NO: 2 in particular were associated with an increased risk of osteoarthritis: 23236, 32358, 47261, 48771, 55292, 60688, 72274, 74209, 77589, 79135, 79456, 79609, 80119, 80766, 81110, 82433, 84121, 84147, 85678, 86699, 86832, 87140 and 88589, where specific embodiments are directed to a polymorphic variant at position 32358, 47261, 74209 and/or 79456. In particular, the following polymorphic variants in SEQ ID NO: 2 were associated with risk of osteoarthritis: an adenine at position 23236, a cytosine at position 32358, a guanine at position 47261, a guanine at position 48771, a cytosine at position 55292, an adenine at position 60688, a guanine at position 72274, a guanine at position 74209, a cytosine at position 77589, an adenine at position 79135, a thymine at position 79456, an adenine at position 79609, an adenine at position 80119, a cytosine at position 80766, an adenine at position 81110, a cytosine at position 82433, a cytosine at position 84121, a thymine at position 84147, a cytosine at position 85678, a thymine at position 86699, an adenine at position 86832, a guanine at position 87140 and an adenine at position 88589.

[0028] Polymorphic variants in and around the *SNW1* locus were tested for association with osteoarthritis. These include polymorphic variants at positions in SEQ ID NO: 3 selected from the group consisting of 218, 1440, 1442, 2611, 4317, 4724, 4788, 5202, 5780, 5974, 6644, 7430, 7938, 8095, 8183, 8312, 8352, 9348, 9378, 9617, 9727, 9834, 9899, 10211, 10377, 10695, 10729, 10730, 11433, 11951, 12697, 12982, 14419, 14501, 14983, 15280, 15475, 15888, 15976, 16307, 16442, 17255, 18948, 19435, 19753, 20021, 20022, 20503, 20590, 21804, 21919, 21990, 22412, 22536, 23432, 23468, 23772, 24325, 24773, 26274, 27440, 28561, 30071, 31764, 33008, 35310, 35460, 37112, 37285, 37747, 38057, 38859,

38860, 39525, 40216, 40281, 41453, 42091, 42513, 42935, 42985, 43003, 43281, 43716, 43866, 44234, 44596, 44871, 45005, 45282, 47178, 47816, 47887, 48134, 48135, 48276, 48400, 48798, 48803, 49146, 49969, 51059, 51064, 53285, 54560, 54748, 54785, 55102, 55644, 55705, 55841, 56623, 56825, 56827, 56892, 59150, 59958, 60231, 60524, 61871, 62226, 63230, 63468, 63787, 65732, 65989, 68832, 69904, 70365, 70886, 73088, 73103, 75934, 75966, 76273, 77943, 78466, 78861, 78872, 79836, 80908, 81509, 83576, 83662, 83782, 84282, 84444, 85129, 85151, 85296, 85809, 86387, 86494, 89786, 89894, 90122, 92067, 92187, 92312, 92824, 93733, 96553 and 96941. Polymorphic variants at the following positions in SEQ ID NO: 3 in particular were associated with an increased risk of osteoarthritis: 4788, 8312, 9378, 9727, 9899, 10211, 27440, 40216, 40281, 42091, 43866, 48803, 51059, 55644, 56623, 73103, 78872, 79836, 85129, 92824 and 96941. In particular, the following polymorphic variants in SEQ ID NO: 3 were associated with risk of osteoarthritis: a guanine at position 4788, a thymine at position 8312, a deletion at position 9378, a cytosine at position 9727, a guanine at position 9899, a cytosine at position 10211, a guanine at position 27440, a guanine at position 40216, a cytosine at position 40281, an adenine at position 42091, a guanine at position 43866, an adenine at position 48803, an adenine at position 51059, an adenine at position 55644, a cytosine at position 56623, a cytosine at position 73103, an adenine at position 78872, a guanine at position 79836, a cytosine at position 85129, a guanine at position 92824 and an adenine at position 96941.

[0029] Based in part upon analyses summarized in Figures 1A-1C, regions with significant association have been identified in regions associated with osteoarthritis. Any polymorphic variants associated with osteoarthritis in a region of significant association can be utilized for embodiments described herein. For example, polymorphic variants in a region spanning chromosome positions 31118000 to 31129000 (approximately 11,000 nucleotides in length) in a *KIAA0296* locus, a region spanning chromosome positions 36914000 to 36931000 (approximately 17,000 nucleotides in length) in a *chrom 4* locus and a region spanning chromosome positions 76196500 to 76221500 (approximately 25,000 nucleotides in length) in a *SNWI* locus have significant association (chromosome positions are within NCBI's Genome build 34).

Additional Polymorphic Variants Associated with Osteoarthritis

[0030] Also provided is a method for identifying polymorphic variants proximal to an incident, founder polymorphic variant associated with osteoarthritis. Thus, featured herein are methods for identifying a polymorphic variation associated with osteoarthritis that is proximal to an incident polymorphic variation associated with osteoarthritis, which comprises identifying a polymorphic variant proximal to the incident polymorphic variant associated with osteoarthritis, where the incident polymorphic variant is in a *KIAA0296*, *chrom 4* or *SNWI* nucleotide sequence or other nucleotide sequence referenced in Table A. The nucleotide sequence often comprises a polynucleotide sequence

selected from the group consisting of (a) a polynucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; (b) a polynucleotide sequence that encodes a polypeptide having an amino acid sequence encoded by a polynucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; and (c) a polynucleotide sequence that encodes a polypeptide having an amino acid sequence that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A or a polynucleotide sequence 90% or more identical to the polynucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A. The presence or absence of an association of the proximal polymorphic variant with osteoarthritis then is determined using a known association method, such as a method described in the Examples hereafter. In an embodiment, the incident polymorphic variant is a polymorphic variant associated with osteoarthritis described herein. In another embodiment, the proximal polymorphic variant identified sometimes is a publicly disclosed polymorphic variant, which for example, sometimes is published in a publicly available database. In other embodiments, the polymorphic variant identified is not publicly disclosed and is discovered using a known method, including, but not limited to, sequencing a region surrounding the incident polymorphic variant in a group of nucleic samples. Thus, multiple polymorphic variants proximal to an incident polymorphic variant are associated with osteoarthritis using this method.

[0031] The proximal polymorphic variant often is identified in a region surrounding the incident polymorphic variant. In certain embodiments, this surrounding region is about 50 kb flanking the first polymorphic variant (*e.g.* about 50 kb 5' of the first polymorphic variant and about 50 kb 3' of the first polymorphic variant), and the region sometimes is composed of shorter flanking sequences, such as flanking sequences of about 40 kb, about 30 kb, about 25 kb, about 20 kb, about 15 kb, about 10 kb, about 7 kb, about 5 kb, or about 2 kb 5' and 3' of the incident polymorphic variant. In other embodiments, the region is composed of longer flanking sequences, such as flanking sequences of about 55 kb, about 60 kb, about 65 kb, about 70 kb, about 75 kb, about 80 kb, about 85 kb, about 90 kb, about 95 kb, or about 100 kb 5' and 3' of the incident polymorphic variant.

[0032] In certain embodiments, polymorphic variants associated with osteoarthritis are identified iteratively. For example, a first proximal polymorphic variant is associated with osteoarthritis using the methods described above and then another polymorphic variant proximal to the first proximal polymorphic variant is identified (*e.g.*, publicly disclosed or discovered) and the presence or absence of an association of one or more other polymorphic variants proximal to the first proximal polymorphic variant with osteoarthritis is determined.

[0033] The methods described herein are useful for identifying or discovering additional polymorphic variants that may be used to further characterize a gene, region or loci associated with a condition, a disease (*e.g.*, osteoarthritis), or a disorder. For example, allelotyping or genotyping data from the additional polymorphic variants may be used to identify a functional mutation or a region of

linkage disequilibrium. In certain embodiments, polymorphic variants identified or discovered within a region comprising the first polymorphic variant associated with osteoarthritis are genotyped using the genetic methods and sample selection techniques described herein, and it can be determined whether those polymorphic variants are in linkage disequilibrium with the first polymorphic variant. The size of the region in linkage disequilibrium with the first polymorphic variant also can be assessed using these genotyping methods. Thus, provided herein are methods for determining whether a polymorphic variant is in linkage disequilibrium with a first polymorphic variant associated with osteoarthritis, and such information can be used in prognosis/diagnosis methods described herein.

Isolated Nucleic Acids

[0034] Featured herein are isolated *KIAA0296*, *chrom 4* or *SNW1* nucleic acid variants depicted in SEQ ID NO: 1-3 or referenced in Table A, and substantially identical nucleic acids thereof. A nucleic acid variant may be represented on one or both strands in a double-stranded nucleic acid or on one chromosomal complement (heterozygous) or both chromosomal complements (homozygous).

[0035] As used herein, the term “nucleic acid” includes DNA molecules (*e.g.*, a complementary DNA (cDNA) and genomic DNA (gDNA)) and RNA molecules (*e.g.*, mRNA, rRNA, siRNA and tRNA) and analogs of DNA or RNA, for example, by use of nucleotide analogs. The nucleic acid molecule can be single-stranded and it is often double-stranded. The term “isolated or purified nucleic acid” refers to nucleic acids that are separated from other nucleic acids present in the natural source of the nucleic acid. For example, with regard to genomic DNA, the term “isolated” includes nucleic acids which are separated from the chromosome with which the genomic DNA is naturally associated. An “isolated” nucleic acid is often free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5’ and/or 3’ ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of 5’ and/or 3’ nucleotide sequences which flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an “isolated” nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. As used herein, the term “gene” refers to a nucleotide sequence that encodes a polypeptide.

[0036] Also included herein are nucleic acid fragments. These fragments often have a nucleotide sequence identical to a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A, a nucleotide sequence substantially identical to a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A, or a nucleotide sequence that is complementary to the foregoing. The nucleic acid fragment may be identical, substantially identical or homologous to a nucleotide sequence in an exon or an intron in a

nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A, and may encode a domain or part of a domain of a polypeptide. Sometimes, the fragment will comprises one or more of the polymorphic variations described herein as being associated with osteoarthritis. The nucleic acid fragment is often 50, 100, or 200 or fewer base pairs in length, and is sometimes about 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 2000, 3000, 4000, 5000, 10000, 15000, or 20000 base pairs in length. A nucleic acid fragment that is complementary to a nucleotide sequence identical or substantially identical to a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A and hybridizes to such a nucleotide sequence under stringent conditions is often referred to as a "probe." Nucleic acid fragments often include one or more polymorphic sites, or sometimes have an end that is adjacent to a polymorphic site as described hereafter.

[0037] An example of a nucleic acid fragment is an oligonucleotide. As used herein, the term "oligonucleotide" refers to a nucleic acid comprising about 8 to about 50 covalently linked nucleotides, often comprising from about 8 to about 35 nucleotides, and more often from about 10 to about 25 nucleotides. The backbone and nucleotides within an oligonucleotide may be the same as those of naturally occurring nucleic acids, or analogs or derivatives of naturally occurring nucleic acids, provided that oligonucleotides having such analogs or derivatives retain the ability to hybridize specifically to a nucleic acid comprising a targeted polymorphism. Oligonucleotides described herein may be used as hybridization probes or as components of prognostic or diagnostic assays, for example, as described herein.

[0038] Oligonucleotides are typically synthesized using standard methods and equipment, such as the ABI™3900 High Throughput DNA Synthesizer and the EXPEDITE™ 8909 Nucleic Acid Synthesizer, both of which are available from Applied Biosystems (Foster City, CA). Analogs and derivatives are exemplified in U.S. Pat. Nos. 4,469,863; 5,536,821; 5,541,306; 5,637,683; 5,637,684; 5,700,922; 5,717,083; 5,719,262; 5,739,308; 5,773,601; 5,886,165; 5,929,226; 5,977,296; 6,140,482; WO 00/56746; WO 01/14398, and related publications. Methods for synthesizing oligonucleotides comprising such analogs or derivatives are disclosed, for example, in the patent publications cited above and in U.S. Pat. Nos. 5,614,622; 5,739,314; 5,955,599; 5,962,674; 6,117,992; in WO 00/75372; and in related publications.

[0039] Oligonucleotides may also be linked to a second moiety. The second moiety may be an additional nucleotide sequence such as a tail sequence (*e.g.*, a polyadenosine tail), an adapter sequence (*e.g.*, phage M13 universal tail sequence), and others. Alternatively, the second moiety may be a non-nucleotide moiety such as a moiety which facilitates linkage to a solid support or a label to facilitate detection of the oligonucleotide. Such labels include, without limitation, a radioactive label, a fluorescent label, a chemiluminescent label, a paramagnetic label, and the like. The second moiety may

be attached to any position of the oligonucleotide, provided the oligonucleotide can hybridize to the nucleic acid comprising the polymorphism.

Uses for Nucleic Acid Sequence

[0040] Nucleic acid coding sequences may be used for diagnostic purposes for detection and control of polypeptide expression. Also, included herein are oligonucleotide sequences such as antisense RNA, small-interfering RNA (siRNA) and DNA molecules and ribozymes that function to inhibit translation of a polypeptide. Antisense techniques and RNA interference techniques are known in the art and are described herein.

[0041] Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. The mechanism of ribozyme action involves sequence specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, hammerhead motif ribozyme molecules may be engineered that specifically and efficiently catalyze endonucleolytic cleavage of RNA sequences corresponding to or complementary to *KIAA0296, chrom 4* or *SNW1* nucleotide sequences or other nucleotide sequences referenced in Table A. Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences, GUA, GUU and GUC. Once identified, short RNA sequences of between fifteen (15) and twenty (20) ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for predicted structural features such as secondary structure that may render the oligonucleotide sequence unsuitable. The suitability of candidate targets may also be evaluated by testing their accessibility to hybridization with complementary oligonucleotides, using ribonuclease protection assays.

[0042] Antisense RNA and DNA molecules, siRNA and ribozymes may be prepared by any method known in the art for the synthesis of RNA molecules. These include techniques for chemically synthesizing oligodeoxyribonucleotides well known in the art such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors which incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

[0043] DNA encoding a polypeptide also may have a number of uses for the diagnosis of diseases, including osteoarthritis, resulting from aberrant expression of a target gene described herein. For example, the nucleic acid sequence may be used in hybridization assays of biopsies or autopsies to diagnose abnormalities of expression or function (e.g., Southern or Northern blot analysis, in situ hybridization assays).

[0044] In addition, the expression of a polypeptide during embryonic development may also be determined using nucleic acid encoding the polypeptide. As addressed, *infra*, production of functionally impaired polypeptide is the cause of various disease states, such as osteoarthritis. *In situ* hybridizations using polypeptide as a probe may be employed to predict problems related to osteoarthritis. Further, as indicated, *infra*, administration of human active polypeptide, recombinantly produced as described herein, may be used to treat disease states related to functionally impaired polypeptide. Alternatively, gene therapy approaches may be employed to remedy deficiencies of functional polypeptide or to replace or compete with dysfunctional polypeptide.

Expression Vectors, Host Cells, and Genetically Engineered Cells

[0045] Provided herein are nucleic acid vectors, often expression vectors, which contain a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A, or a substantially identical sequence thereof. As used herein, the term “vector” refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked and can include a plasmid, cosmid, or viral vector. The vector can be capable of autonomous replication or it can integrate into a host DNA. Viral vectors may include replication defective retroviruses, adenoviruses and adeno-associated viruses for example.

[0046] A vector can include a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A in a form suitable for expression of an encoded target polypeptide or target nucleic acid in a host cell. A “target polypeptide” is a polypeptide encoded by a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A, or a substantially identical nucleotide sequence thereof. The recombinant expression vector typically includes one or more regulatory sequences operatively linked to the nucleic acid sequence to be expressed. The term “regulatory sequence” includes promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence, as well as tissue-specific regulatory and/or inducible sequences. The design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of polypeptide desired, and the like. Expression vectors can be introduced into host cells to produce target polypeptides, including fusion polypeptides.

[0047] Recombinant expression vectors can be designed for expression of target polypeptides in prokaryotic or eukaryotic cells. For example, target polypeptides can be expressed in *E. coli*, insect cells (*e.g.*, using baculovirus expression vectors), yeast cells, or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology 185*, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

[0048] Expression of polypeptides in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion polypeptides. Fusion vectors add a number of amino acids to a polypeptide encoded therein, usually to the amino terminus of the recombinant polypeptide. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant polypeptide; 2) to increase the solubility of the recombinant polypeptide; and 3) to aid in the purification of the recombinant polypeptide by acting as a ligand in affinity purification. Often, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant polypeptide to enable separation of the recombinant polypeptide from the fusion moiety subsequent to purification of the fusion polypeptide. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith & Johnson, *Gene* 67: 31-40 (1988)), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding polypeptide, or polypeptide A, respectively, to the target recombinant polypeptide.

[0049] Purified fusion polypeptides can be used in screening assays and to generate antibodies specific for target polypeptides. In a therapeutic embodiment, fusion polypeptide expressed in a retroviral expression vector is used to infect bone marrow cells that are subsequently transplanted into irradiated recipients. The pathology of the subject recipient is then examined after sufficient time has passed (*e.g.*, six (6) weeks).

[0050] Expressing the polypeptide in host bacteria with an impaired capacity to proteolytically cleave the recombinant polypeptide is often used to maximize recombinant polypeptide expression (Gottesman, S., *Gene Expression Technology: Methods in Enzymology*, Academic Press, San Diego, California 185: 119-128 (1990)). Another strategy is to alter the nucleotide sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, *Nucleic Acids Res.* 20: 2111-2118 (1992)). Such alteration of nucleotide sequences can be carried out by standard DNA synthesis techniques.

[0051] When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. Recombinant mammalian expression vectors are often capable of directing expression of the nucleic acid in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Non-limiting examples of suitable tissue-specific promoters include an albumin promoter (liver-specific; Pinkert *et al.*, *Genes Dev.* 1: 268-277 (1987)), lymphoid-specific promoters (Calame & Eaton, *Adv. Immunol.* 43: 235-275 (1988)), promoters of T cell receptors (Winoto & Baltimore, *EMBO J.* 8: 729-733 (1989)) promoters of immunoglobulins (Banerji *et al.*, *Cell* 33: 729-740 (1983); Queen & Baltimore, *Cell* 33: 741-748 (1983)), neuron-specific

promoters (*e.g.*, the neurofilament promoter; Byrne & Ruddle, *Proc. Natl. Acad. Sci. USA* 86: 5473-5477 (1989)), pancreas-specific promoters (Edlund *et al.*, *Science* 230: 912-916 (1985)), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are sometimes utilized, for example, the murine hox promoters (Kessel & Gruss, *Science* 249: 374-379 (1990)) and the α -fetopolypeptide promoter (Campes & Tilghman, *Genes Dev.* 3: 537-546 (1989)).

[0052] A *KIAA0296*, *chrom 4* or *SNW1* nucleic acid or other nucleic acid referenced in Table A also may be cloned into an expression vector in an antisense orientation. Regulatory sequences (*e.g.*, viral promoters and/or enhancers) operatively linked to a *KIAA0296*, *chrom 4* or *SNW1* nucleic acid or other nucleic acid referenced in Table A cloned in the antisense orientation can be chosen for directing constitutive, tissue specific or cell type specific expression of antisense RNA in a variety of cell types. Antisense expression vectors can be in the form of a recombinant plasmid, phagemid or attenuated virus. For a discussion of the regulation of gene expression using antisense genes *see, e.g.*, Weintraub *et al.*, *Antisense RNA as a molecular tool for genetic analysis, Reviews - Trends in Genetics*, Vol. 1(1) (1986).

[0053] Also provided herein are host cells that include a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A within a recombinant expression vector or a fragment of such a nucleotide sequence which facilitate homologous recombination into a specific site of the host cell genome. The terms "host cell" and "recombinant host cell" are used interchangeably herein. Such terms refer not only to the particular subject cell but rather also to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein. A host cell can be any prokaryotic or eukaryotic cell. For example, a target polypeptide can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

[0054] Vectors can be introduced into host cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, transduction/infection, DEAE-dextran-mediated transfection, lipofection, or electroporation.

[0055] A host cell provided herein can be used to produce (*i.e.*, express) a target polypeptide or a substantially identical polypeptide thereof. Accordingly, further provided are methods for producing a target polypeptide using host cells described herein. In one embodiment, the method includes culturing host cells into which a recombinant expression vector encoding a target polypeptide has been introduced

in a suitable medium such that a target polypeptide is produced. In another embodiment, the method further includes isolating a target polypeptide from the medium or the host cell.

[0056] Also provided are cells or purified preparations of cells which include a *KIAA0296, chrom 4* or *SNWI* transgene, or other transgene in Table A, or which otherwise misexpress target polypeptide. Cell preparations can consist of human or non-human cells, *e.g.*, rodent cells, *e.g.*, mouse or rat cells, rabbit cells, or pig cells. In preferred embodiments, the cell or cells include a *KIAA0296, chrom 4* or *SNWI* transgene or other transgene referenced in Table A (*e.g.*, a heterologous form of a *KIAA0296, chrom 4* or *SNWI* gene or other gene referenced in Table A, such as a human gene expressed in non-human cells). The transgene can be misexpressed, *e.g.*, overexpressed or underexpressed. In other preferred embodiments, the cell or cells include a gene which misexpress an endogenous target polypeptide (*e.g.*, expression of a gene is disrupted, also known as a knockout). Such cells can serve as a model for studying disorders which are related to mutated or mis-expressed alleles or for use in drug screening. Also provided are human cells (*e.g.*, a hematopoietic stem cells) transfected with a *KIAA0296, chrom 4* or *SNWI* nucleic acid or other nucleic acid referenced in Table A.

[0057] Also provided are cells or a purified preparation thereof (*e.g.*, human cells) in which an endogenous *KIAA0296, chrom 4* or *SNWI* nucleic acid or other nucleic acid referenced in Table A is under the control of a regulatory sequence that does not normally control the expression of the endogenous gene. The expression characteristics of an endogenous gene within a cell (*e.g.*, a cell line or microorganism) can be modified by inserting a heterologous DNA regulatory element into the genome of the cell such that the inserted regulatory element is operably linked to the corresponding endogenous gene. For example, an endogenous corresponding gene (*e.g.*, a gene which is “transcriptionally silent,” not normally expressed, or expressed only at very low levels) may be activated by inserting a regulatory element which is capable of promoting the expression of a normally expressed gene product in that cell. Techniques such as targeted homologous recombinations, can be used to insert the heterologous DNA as described in, *e.g.*, Chappel, US 5,272,071; WO 91/06667, published on May 16, 1991.

Transgenic Animals

[0058] Non-human transgenic animals that express a heterologous target polypeptide (*e.g.*, expressed from a *KIAA0296, chrom 4* or *SNWI* nucleic acid or other nucleic acid referenced in Table A, or substantially identical sequence thereof) can be generated. Such animals are useful for studying the function and/or activity of a target polypeptide and for identifying and/or evaluating modulators of the activity of *KIAA0296, chrom 4* or *SNWI* nucleic acids, other nucleic acids referenced in Table A, and encoded polypeptides. As used herein, a “transgenic animal” is a non-human animal such as a mammal (*e.g.*, a non-human primate such as chimpanzee, baboon, or macaque; an ungulate such as an equine, bovine, or caprine; or a rodent such as a rat, a mouse, or an Israeli sand rat), a bird (*e.g.*, a chicken or a

turkey), an amphibian (e.g., a frog, salamander, or newt), or an insect (e.g., *Drosophila melanogaster*), in which one or more of the cells of the animal includes a transgene. A transgene is exogenous DNA or a rearrangement (e.g., a deletion of endogenous chromosomal DNA) that is often integrated into or occurs in the genome of cells in a transgenic animal. A transgene can direct expression of an encoded gene product in one or more cell types or tissues of the transgenic animal, and other transgenes can reduce expression (e.g., a knockout). Thus, a transgenic animal can be one in which an endogenous nucleic acid homologous to a *KIAA0296*, *chrom 4* or *SNW1* nucleic acid or other nucleic acid referenced in Table A has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal (e.g., an embryonic cell of the animal) prior to development of the animal.

[0059] Intronic sequences and polyadenylation signals can also be included in the transgene to increase expression efficiency of the transgene. One or more tissue-specific regulatory sequences can be operably linked to a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A to direct expression of an encoded polypeptide to particular cells. A transgenic founder animal can be identified based upon the presence of a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A in its genome and/or expression of encoded mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A can further be bred to other transgenic animals carrying other transgenes.

[0060] Target polypeptides can be expressed in transgenic animals or plants by introducing, for example, a *KIAA0296*, *chrom 4* or *SNW1* nucleic acid or other nucleic acid referenced in Table A into the genome of an animal that encodes the target polypeptide. In preferred embodiments the nucleic acid is placed under the control of a tissue specific promoter, e.g., a milk or egg specific promoter, and recovered from the milk or eggs produced by the animal. Also included is a population of cells from a transgenic animal.

Target Polypeptides

[0061] Also featured herein are isolated target polypeptides, which are encoded by a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or a nucleotide sequence referenced in Table A (e.g., SEQ ID NO: 1-3 or a sequence referenced in Table A), or a substantially identical nucleotide sequence thereof. The term "polypeptide" as used herein includes proteins and peptides. An "isolated" or "purified" polypeptide or protein is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. In one embodiment, the language "substantially free"

means preparation of a target polypeptide having less than about 30%, 20%, 10% and more preferably 5% (by dry weight), of non-target polypeptide (also referred to herein as a “contaminating protein”), or of chemical precursors or non-target chemicals. When the target polypeptide or a biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, specifically, where culture medium represents less than about 20%, sometimes less than about 10%, and often less than about 5% of the volume of the polypeptide preparation. Isolated or purified target polypeptide preparations are sometimes 0.01 milligrams or more or 0.1 milligrams or more, and often 1.0 milligrams or more and 10 milligrams or more in dry weight.

[0062] Further included herein are target polypeptide fragments. The polypeptide fragment may be a domain or part of a domain of a target polypeptide. The polypeptide fragment may have increased, decreased or unexpected biological activity. The polypeptide fragment is often 50 or fewer, 100 or fewer, or 200 or fewer amino acids in length, and is sometimes 300, 400, 500, 600, 700, or 900 or fewer amino acids in length. Specific embodiments are directed to a *PTPNI* polypeptide fragment (e.g., rs2282146 in Table A), such as a catalytic domain starting at about amino acid 3 and ending at about amino acid 279. Other embodiments are directed to a *KCNS1* polypeptide fragment (e.g., rs734784 in Table A), such as a voltage gated postassium ion channel domain (e.g., starting at about amino acid 21 and ending at about amino acid 509), a postassium channel tetramerization domain (e.g., starting at about amino acid 52 and ending at about amino acid 155) or an ion transport protein domain (e.g., starting at about amino acid 271 and ending at about amino acid 456), for example. Certain embodiments are directed to a *ANXA6* polypeptide fragment (e.g., rs1012414 in Table A), such as an annexin domain starting at about amino acid 5 and ending at about amino acid 325, an annexin domain starting at about amino acid 179 and ending at about amino acid 507, or an annexin domain starting at about amino acid 355 and ending at about amino acid 673 in isoform 1 or isoform 2 (e.g., an isoform 1 sequence can be accessed using accession number NP_001146 and an isoform 2 sequence can be accessed using accession number NP_004024; isoform 2 lacks exon 21 and encodes a protein isoform lacking the six amino acids VAAEIL). Amino acid sequences can be accessed using information in Table A.

[0063] Substantially identical target polypeptides may depart from the amino acid sequences of target polypeptides in different manners. For example, conservative amino acid modifications may be introduced at one or more positions in the amino acid sequences of target polypeptides. A “conservative amino acid substitution” is one in which the amino acid is replaced by another amino acid having a similar structure and/or chemical function. Families of amino acid residues having similar structures and functions are well known. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side

chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Also, essential and non-essential amino acids may be replaced. A “non-essential” amino acid is one that can be altered without abolishing or substantially altering the biological function of a target polypeptide, whereas altering an “essential” amino acid abolishes or substantially alters the biological function of a target polypeptide. Amino acids that are conserved among target polypeptides are typically essential amino acids. In certain embodiments, the polypeptide includes one or more non-synonymous polymorphic variants associated with osteoarthritis, as described above (e.g., a valine encoded by rs734784, a valine encoded by rs1042164, a glutamate encoded by rs749670, a threonine encoded by rs955592 and a glycine encoded by rs1040461).

[0064] Also, target polypeptides may exist as chimeric or fusion polypeptides. As used herein, a target “chimeric polypeptide” or target “fusion polypeptide” includes a target polypeptide linked to a non-target polypeptide. A “non-target polypeptide” refers to a polypeptide having an amino acid sequence corresponding to a polypeptide which is not substantially identical to the target polypeptide, which includes, for example, a polypeptide that is different from the target polypeptide and derived from the same or a different organism. The target polypeptide in the fusion polypeptide can correspond to an entire or nearly entire target polypeptide or a fragment thereof. The non-target polypeptide can be fused to the N-terminus or C-terminus of the target polypeptide.

[0065] Fusion polypeptides can include a moiety having high affinity for a ligand. For example, the fusion polypeptide can be a GST-target fusion polypeptide in which the target sequences are fused to the C-terminus of the GST sequences, or a polyhistidine-target fusion polypeptide in which the target polypeptide is fused at the N- or C-terminus to a string of histidine residues. Such fusion polypeptides can facilitate purification of recombinant target polypeptide. Expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide), and a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A, or a substantially identical nucleotide sequence thereof, can be cloned into an expression vector such that the fusion moiety is linked in-frame to the target polypeptide. Further, the fusion polypeptide can be a target polypeptide containing a heterologous signal sequence at its N-terminus. In certain host cells (e.g., mammalian host cells), expression, secretion, cellular internalization, and cellular localization of a target polypeptide can be increased through use of a heterologous signal sequence. Fusion polypeptides can also include all or a part of a serum polypeptide (e.g., an IgG constant region or human serum albumin).

[0066] Target polypeptides can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. Administration of these target polypeptides can be used to affect the bioavailability of a substrate of the target polypeptide and may effectively increase target polypeptide biological activity in a cell. Target fusion polypeptides may be useful therapeutically for the treatment of disorders caused by, for example, (i) aberrant modification or mutation of a gene encoding a target polypeptide; (ii) mis-

regulation of the gene encoding the target polypeptide; and (iii) aberrant post-translational modification of a target polypeptide. Also, target polypeptides can be used as immunogens to produce anti-target antibodies in a subject, to purify target polypeptide ligands or binding partners, and in screening assays to identify molecules which inhibit or enhance the interaction of a target polypeptide with a substrate.

[0067] In addition, polypeptides can be chemically synthesized using techniques known in the art (See, *e.g.*, Creighton, 1983 Proteins. New York, N.Y.: W. H. Freeman and Company; and Hunkapiller et al., (1984) Nature July 12 -18;310(5973):105-11). For example, a relative short fragment can be synthesized by use of a peptide synthesizer. Furthermore, if desired, non-classical amino acids or chemical amino acid analogs can be introduced as a substitution or addition into the fragment sequence. Non-classical amino acids include, but are not limited to, to the D-isomers of the common amino acids, 2,4-diaminobutyric acid, α -amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid, γ -Abu, ϵ -Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, homocitrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, β -alanine, fluoroamino acids, designer amino acids such as β -methyl amino acids, Ca-methyl amino acids, Na-methyl amino acids, and amino acid analogs in general. Furthermore, the amino acid can be D (dextrorotary) or L (levorotary).

[0068] Polypeptides and polypeptide fragments sometimes are differentially modified during or after translation, *e.g.*, by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications may be carried out by known techniques, including but not limited, to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, NaBH₄; acetylation, formylation, oxidation, reduction; metabolic synthesis in the presence of tunicamycin; and the like. Additional post-translational modifications include, for example, N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends), attachment of chemical moieties to the amino acid backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and addition or deletion of an N-terminal methionine residue as a result of prokaryotic host cell expression. The polypeptide fragments may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the polypeptide.

[0069] Also provided are chemically modified derivatives of polypeptides that can provide additional advantages such as increased solubility, stability and circulating time of the polypeptide, or decreased immunogenicity (*see e.g.*, U.S. Pat. No: 4,179,337. The chemical moieties for derivitization may be selected from water soluble polymers such as polyethylene glycol, ethylene glycol/propylene glycol copolymers, carboxymethylcellulose, dextran, polyvinyl alcohol and the like. The polypeptides may be modified at random positions within the molecule, or at predetermined positions within the molecule and may include one, two, three or more attached chemical moieties.

[0070] The polymer may be of any molecular weight, and may be branched or unbranched. For polyethylene glycol, the preferred molecular weight is between about 1 kDa and about 100 kDa (the term “about” indicating that in preparations of polyethylene glycol, some molecules will weigh more, some less, than the stated molecular weight) for ease in handling and manufacturing. Other sizes may be used, depending on the desired therapeutic profile (e.g., the duration of sustained release desired, the effects, if any on biological activity, the ease in handling, the degree or lack of antigenicity and other known effects of the polyethylene glycol to a therapeutic protein or analog).

[0071] The polymers should be attached to the polypeptide with consideration of effects on functional or antigenic domains of the polypeptide. There are a number of attachment methods available to those skilled in the art (e.g., EP 0 401 384 (coupling PEG to G-CSF) and Malik et al. (1992) Exp Hematol. September;20(8):1028-35 (pegylation of GM-CSF using tresyl chloride)). For example, polyethylene glycol may be covalently bound through amino acid residues via a reactive group, such as a free amino or carboxyl group. Reactive groups are those to which an activated polyethylene glycol molecule may be bound. The amino acid residues having a free amino group may include lysine residues and the N-terminal amino acid residues; those having a free carboxyl group may include aspartic acid residues, glutamic acid residues and the C-terminal amino acid residue. Sulfhydryl groups may also be used as a reactive group for attaching the polyethylene glycol molecules. For therapeutic purposes, the attachment sometimes is at an amino group, such as attachment at the N-terminus or lysine group.

[0072] Proteins can be chemically modified at the N-terminus. Using polyethylene glycol as an illustration of such a composition, one may select from a variety of polyethylene glycol molecules (by molecular weight, branching, and the like), the proportion of polyethylene glycol molecules to protein (polypeptide) molecules in the reaction mix, the type of pegylation reaction to be performed, and the method of obtaining the selected N-terminally pegylated protein. The method of obtaining the N-terminally pegylated preparation (i.e., separating this moiety from other monopegylated moieties if necessary) may be by purification of the N-terminally pegylated material from a population of pegylated protein molecules. Selective proteins chemically modified at the N-terminus may be accomplished by reductive alkylation, which exploits differential reactivity of different types of primary amino groups (lysine versus the N-terminal) available for derivatization in a particular protein. Under the appropriate reaction conditions, substantially selective derivatization of the protein at the N-terminus with a carbonyl group containing polymer is achieved.

Substantially Identical Nucleic Acids and Polypeptides

[0073] Nucleotide sequences and polypeptide sequences that are substantially identical to a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A and the target polypeptide sequences encoded by those nucleotide sequences, respectively, are included

herein. The term “substantially identical” as used herein refers to two or more nucleic acids or polypeptides sharing one or more identical nucleotide sequences or polypeptide sequences, respectively. Included are nucleotide sequences or polypeptide sequences that are 55% or more, 60% or more, 65% or more, 70% or more, 75% or more, 80% or more, 85% or more, 90% or more, 95% or more (each often within a 1%, 2%, 3% or 4% variability) identical to a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence, or other nucleotide sequence referenced in Table A, or the encoded target polypeptide amino acid sequences. One test for determining whether two nucleic acids are substantially identical is to determine the percent of identical nucleotide sequences or polypeptide sequences shared between the nucleic acids or polypeptides.

[0074] Calculations of sequence identity are often performed as follows. Sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). The length of a reference sequence aligned for comparison purposes is sometimes 30% or more, 40% or more, 50% or more, often 60% or more, and more often 70% or more, 80% or more, 90% or more, or 100% of the length of the reference sequence. The nucleotides or amino acids at corresponding nucleotide or polypeptide positions, respectively, are then compared among the two sequences. When a position in the first sequence is occupied by the same nucleotide or amino acid as the corresponding position in the second sequence, the nucleotides or amino acids are deemed to be identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, introduced for optimal alignment of the two sequences.

[0075] Comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. Percent identity between two amino acid or nucleotide sequences can be determined using the algorithm of Meyers & Miller, *CABIOS* 4: 11-17 (1989), which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4. Also, percent identity between two amino acid sequences can be determined using the Needleman & Wunsch, *J. Mol. Biol.* 48: 444-453 (1970) algorithm which has been incorporated into the GAP program in the GCG software package (available at the http address www.gcg.com), using either a Blossum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. Percent identity between two nucleotide sequences can be determined using the GAP program in the GCG software package (available at http address www.gcg.com), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. A set of parameters often used is a Blossum 62 scoring matrix with a gap open penalty of 12, a gap extend penalty of 4, and a frameshift gap penalty of 5.

[0076] Another manner for determining if two nucleic acids are substantially identical is to assess whether a polynucleotide homologous to one nucleic acid will hybridize to the other nucleic acid under stringent conditions. As use herein, the term “stringent conditions” refers to conditions for hybridization and washing. Stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. , 6.3.1-6.3.6 (1989). Aqueous and non-aqueous methods are described in that reference and either can be used. An example of stringent hybridization conditions is hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 50°C. Another example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 55°C. A further example of stringent hybridization conditions is hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 60°C. Often, stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 65°C. More often, stringency conditions are 0.5M sodium phosphate, 7% SDS at 65°C, followed by one or more washes at 0.2X SSC, 1% SDS at 65°C.

[0077] An example of a substantially identical nucleotide sequence to a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A is one that has a different nucleotide sequence but still encodes the same polypeptide sequence encoded by the nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A. Another example is a nucleotide sequence that encodes a polypeptide having a polypeptide sequence that is more than 70% or more identical to, sometimes more than 75% or more, 80% or more, or 85% or more identical to, and often more than 90% or more and 95% or more identical to a polypeptide sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A.

[0078] Nucleotide sequences in SEQ ID NO: 1-3 or referenced in Table A and amino acid sequences of encoded polypeptides can be used as “query sequences” to perform a search against public databases to identify other family members or related sequences, for example. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul *et al.*, *J. Mol. Biol.* 215: 403-10 (1990). BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to nucleotide sequences in SEQ ID NO: 1-3 or referenced in Table A. BLAST polypeptide searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to polypeptides encoded by the nucleotide sequences of SEQ ID NO: 1-3 or referenced in Table A. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.*, *Nucleic Acids Res.* 25(17): 3389-3402 (1997). When utilizing BLAST and Gapped BLAST programs, default parameters of

the respective programs (*e.g.*, XBLAST and NBLAST) can be used (*see* the http address www.ncbi.nlm.nih.gov).

[0079] A nucleic acid that is substantially identical to a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A may include polymorphic sites at positions equivalent to those described herein when the sequences are aligned. For example, using the alignment procedures described herein, SNPs in a sequence substantially identical to a sequence in SEQ ID NO: 1-3 or referenced in Table A can be identified at nucleotide positions that match (*i.e.*, align) with nucleotides at SNP positions in each nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A. Also, where a polymorphic variation results in an insertion or deletion, insertion or deletion of a nucleotide sequence from a reference sequence can change the relative positions of other polymorphic sites in the nucleotide sequence.

[0080] Substantially identical nucleotide and polypeptide sequences include those that are naturally occurring, such as allelic variants (same locus), splice variants, homologs (different locus), and orthologs (different organism) or can be non-naturally occurring. Non-naturally occurring variants can be generated by mutagenesis techniques, including those applied to polynucleotides, cells, or organisms. The variants can contain nucleotide substitutions, deletions, inversions and insertions. Variation can occur in either or both the coding and non-coding regions. The variations can produce both conservative and non-conservative amino acid substitutions (as compared in the encoded product). Orthologs, homologs, allelic variants, and splice variants can be identified using methods known in the art. These variants normally comprise a nucleotide sequence encoding a polypeptide that is 50% or more, about 55% or more, often about 70-75% or more or about 80-85% or more, and sometimes about 90-95% or more identical to the amino acid sequences of target polypeptides or a fragment thereof. Such nucleic acid molecules can readily be identified as being able to hybridize under stringent conditions to a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A or a fragment of this sequence. Nucleic acid molecules corresponding to orthologs, homologs, and allelic variants of a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A can further be identified by mapping the sequence to the same chromosome or locus as the nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A.

[0081] Also, substantially identical nucleotide sequences may include codons that are altered with respect to the naturally occurring sequence for enhancing expression of a target polypeptide in a particular expression system. For example, the nucleic acid can be one in which one or more codons are altered, and often 10% or more or 20% or more of the codons are altered for optimized expression in bacteria (*e.g.*, *E. coli*), yeast (*e.g.*, *S. cerevisiae*), human (*e.g.*, 293 cells), insect, or rodent (*e.g.*, hamster) cells.

Methods for Identifying Risk of Osteoarthritis

[0082] Methods for prognosing and diagnosing osteoarthritis are included herein. These methods include detecting the presence or absence of one or more polymorphic variations in a nucleotide sequence associated with osteoarthritis, such as variants in or around the loci set forth herein, or a substantially identical sequence thereof, in a sample from a subject, where the presence of a polymorphic variant described herein is indicative of a risk of osteoarthritis. Determining a risk of osteoarthritis sometimes refers to determining whether an individual is at an increased risk of osteoarthritis (*e.g.*, intermediate risk or higher risk).

[0083] Thus, featured herein is a method for identifying a subject who is at risk of osteoarthritis, which comprises detecting an aberration associated with osteoarthritis in a nucleic acid sample from the subject. An embodiment is a method for detecting a risk of osteoarthritis in a subject, which comprises detecting the presence or absence of a polymorphic variation associated with osteoarthritis at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject, where the nucleotide sequence comprises a polynucleotide sequence selected from the group consisting of: (a) a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; (b) a nucleotide sequence which encodes a polypeptide consisting of an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A, or a nucleotide sequence about 90% or more identical to a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; and (d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising the polymorphic site; whereby the presence of the polymorphic variation is indicative of a predisposition to osteoarthritis in the subject. In certain embodiments, polymorphic variants at the positions described herein are detected for determining a risk of osteoarthritis, and polymorphic variants at positions in linkage disequilibrium with these positions are detected for determining a risk of osteoarthritis. As used herein, the terms "SEQ ID NO: 1-3" and other nucleotide sequences "referenced in Table A" refers to individual sequences in SEQ ID NO: 1, 2 or 3 or any individual sequence referenced in Table A, each sequence being separately applicable to embodiments described herein.

[0084] Risk of osteoarthritis sometimes is expressed as a probability, such as an odds ratio, percentage, or risk factor. Risk often is based upon the presence or absence of one or more polymorphic variants described herein, and also may be based in part upon phenotypic traits of the individual being tested. Methods for calculating risk based upon patient data are well known (*see, e.g.*, Agresti, *Categorical Data Analysis*, 2nd Ed. 2002. Wiley). Allelotyping and genotyping analyses may be carried out in populations other than those exemplified herein to enhance the predictive power of the prognostic method. These further analyses are executed in view of the exemplified procedures described herein, and may be based upon the same polymorphic variations or additional polymorphic variations.

[0085] In certain embodiments, determining the presence of a combination of two or more polymorphic variants associated with osteoarthritis in one or more genetic loci (e.g., one or more genes) of the sample is determined to identify, quantify and/or estimate, risk of osteoarthritis. The risk often is the probability of having or developing osteoarthritis. The risk sometimes is expressed as a relative risk with respect to a population average risk of osteoarthritis, and sometimes is expressed as a relative risk with respect to the lowest risk group. Such relative risk assessments often are based upon penetrance values determined by statistical methods, and are particularly useful to clinicians and insurance companies for assessing risk of osteoarthritis (e.g., a clinician can target appropriate detection, prevention and therapeutic regimens to a patient after determining the patient's risk of osteoarthritis, and an insurance company can fine tune actuarial tables based upon population genotype assessments of osteoarthritis risk). Risk of osteoarthritis sometimes is expressed as an odds ratio, which is the odds of a particular person having a genotype has or will develop osteoarthritis with respect to another genotype group (e.g., the most disease protective genotype or population average). In related embodiments, the determination is utilized to identify a subject at risk of osteoarthritis. In an embodiment, two or more polymorphic variations are detected in two or more regions in human genomic DNA associated with increased risk of osteoarthritis, such as a locus containing a *KIAA0296*, *chrom 4* or *SNW1* or other locus referenced in Table A, for example. In certain embodiments, 3 or more, or 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 or more polymorphic variants are detected in the sample. In specific embodiments, polymorphic variants are detected in a *KIAA0296*, *chrom 4* or *SNW1* region or other region referenced in Table A, for example. In another embodiment, polymorphic variants are detected at two or three positions in a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A. In certain embodiments, polymorphic variants are detected at other genetic loci (e.g., the polymorphic variants can be detected in a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A in addition to other loci or only in other loci), where the other loci include but are not limited to those described in concurrently-filed patent applications having attorney docket number 524593008700, 524593008800, 524593008900, 524593009000, 524593009100 or 524593009200, each of which is incorporated herein by reference in its entirety.

[0086] Results from prognostic tests may be combined with other test results to diagnose osteoarthritis. For example, prognostic results may be gathered, a patient sample may be ordered based on a determined predisposition to osteoarthritis, the patient sample is analyzed, and the results of the analysis may be utilized to diagnose osteoarthritis. Also osteoarthritis diagnostic method can be developed from studies used to generate prognostic methods in which populations are stratified into subpopulations having different progressions of osteoarthritis. In another embodiment, prognostic results may be gathered, a patient's risk factors for developing osteoarthritis (e.g., age, weight, race, diet) analyzed, and a patient sample may be ordered based on a determined predisposition to osteoarthritis.

[0087] The nucleic acid sample typically is isolated from a biological sample obtained from a subject. For example, nucleic acid can be isolated from blood, saliva, sputum, urine, cell scrapings, and biopsy tissue. The nucleic acid sample can be isolated from a biological sample using standard techniques, such as the technique described in Example 2. As used herein, the term “subject” refers primarily to humans but also refers to other mammals such as dogs, cats, and ungulates (*e.g.*, cattle, sheep, and swine). Subjects also include avians (*e.g.*, chickens and turkeys), reptiles, and fish (*e.g.*, salmon), as embodiments described herein can be adapted to nucleic acid samples isolated from any of these organisms. The nucleic acid sample may be isolated from the subject and then directly utilized in a method for determining the presence of a polymorphic variant, or alternatively, the sample may be isolated and then stored (*e.g.*, frozen) for a period of time before being subjected to analysis.

[0088] The presence or absence of a polymorphic variant is determined using one or both chromosomal complements represented in the nucleic acid sample. Determining the presence or absence of a polymorphic variant in both chromosomal complements represented in a nucleic acid sample from a subject having a copy of each chromosome is useful for determining the zygosity of an individual for the polymorphic variant (*i.e.*, whether the individual is homozygous or heterozygous for the polymorphic variant). Any oligonucleotide-based diagnostic may be utilized to determine whether a sample includes the presence or absence of a polymorphic variant in a sample. For example, primer extension methods, ligase sequence determination methods (*e.g.*, U.S. Pat. Nos. 5,679,524 and 5,952,174, and WO 01/27326), mismatch sequence determination methods (*e.g.*, U.S. Pat. Nos. 5,851,770; 5,958,692; 6,110,684; and 6,183,958), microarray sequence determination methods, restriction fragment length polymorphism (RFLP), single strand conformation polymorphism detection (SSCP) (*e.g.*, U.S. Pat. Nos. 5,891,625 and 6,013,499), PCR-based assays (*e.g.*, TAQMAN[®] PCR System (Applied Biosystems)), and nucleotide sequencing methods may be used.

[0089] Oligonucleotide extension methods typically involve providing a pair of oligonucleotide primers in a polymerase chain reaction (PCR) or in other nucleic acid amplification methods for the purpose of amplifying a region from the nucleic acid sample that comprises the polymorphic variation. One oligonucleotide primer is complementary to a region 3' of the polymorphism and the other is complementary to a region 5' of the polymorphism. A PCR primer pair may be used in methods disclosed in U.S. Pat. Nos. 4,683,195; 4,683,202, 4,965,188; 5,656,493; 5,998,143; 6,140,054; WO 01/27327; and WO 01/27329 for example. PCR primer pairs may also be used in any commercially available machines that perform PCR, such as any of the GENEAMP[®] Systems available from Applied Biosystems. Also, those of ordinary skill in the art will be able to design oligonucleotide primers based upon a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A using knowledge available in the art.

[0090] Also provided is an extension oligonucleotide that hybridizes to the amplified fragment adjacent to the polymorphic variation. As used herein, the term “adjacent” refers to the 3’ end of the extension oligonucleotide being often 1 nucleotide from the 5’ end of the polymorphic site, and sometimes 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from the 5’ end of the polymorphic site, in the nucleic acid when the extension oligonucleotide is hybridized to the nucleic acid. The extension oligonucleotide then is extended by one or more nucleotides, and the number and/or type of nucleotides that are added to the extension oligonucleotide determine whether the polymorphic variant is present. Oligonucleotide extension methods are disclosed, for example, in U.S. Pat. Nos. 4,656,127; 4,851,331; 5,679,524; 5,834,189; 5,876,934; 5,908,755; 5,912,118; 5,976,802; 5,981,186; 6,004,744; 6,013,431; 6,017,702; 6,046,005; 6,087,095; 6,210,891; and WO 01/20039. Oligonucleotide extension methods using mass spectrometry are described, for example, in U.S. Pat. Nos. 5,547,835; 5,605,798; 5,691,141; 5,849,542; 5,869,242; 5,928,906; 6,043,031; and 6,194,144, and a method often utilized is described herein in Example 2.

[0091] A microarray can be utilized for determining whether a polymorphic variant is present or absent in a nucleic acid sample. A microarray may include any oligonucleotides described herein, and methods for making and using oligonucleotide microarrays suitable for diagnostic use are disclosed in U.S. Pat. Nos. 5,492,806; 5,525,464; 5,589,330; 5,695,940; 5,849,483; 6,018,041; 6,045,996; 6,136,541; 6,142,681; 6,156,501; 6,197,506; 6,223,127; 6,225,625; 6,229,911; 6,239,273; WO 00/52625; WO 01/25485; and WO 01/29259. The microarray typically comprises a solid support and the oligonucleotides may be linked to this solid support by covalent bonds or by non-covalent interactions. The oligonucleotides may also be linked to the solid support directly or by a spacer molecule. A microarray may comprise one or more oligonucleotides complementary to a polymorphic site set forth herein.

[0092] A kit also may be utilized for determining whether a polymorphic variant is present or absent in a nucleic acid sample. A kit often comprises one or more pairs of oligonucleotide primers useful for amplifying a fragment of a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A or a substantially identical sequence thereof, where the fragment includes a polymorphic site. The kit sometimes comprises a polymerizing agent, for example, a thermostable nucleic acid polymerase such as one disclosed in U.S. Pat. Nos. 4,889,818 or 6,077,664. Also, the kit often comprises an elongation oligonucleotide that hybridizes to a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A in a nucleic acid sample adjacent to the polymorphic site. Where the kit includes an elongation oligonucleotide, it also often comprises chain elongating nucleotides, such as dATP, dTTP, dGTP, dCTP, and dITP, including analogs of dATP, dTTP, dGTP, dCTP and dITP, provided that such analogs are substrates for a thermostable nucleic acid polymerase and can be incorporated into a nucleic acid chain elongated from the extension oligonucleotide. Along with chain

elongating nucleotides would be one or more chain terminating nucleotides such as ddATP, ddTTP, ddGTP, ddCTP, and the like. In an embodiment, the kit comprises one or more oligonucleotide primer pairs, a polymerizing agent, chain elongating nucleotides, at least one elongation oligonucleotide, and one or more chain terminating nucleotides. Kits optionally include buffers, vials, microtiter plates, and instructions for use.

[0093] An individual identified as being at risk of osteoarthritis may be heterozygous or homozygous with respect to the allele associated with a higher risk of osteoarthritis. A subject homozygous for an allele associated with an increased risk of osteoarthritis is at a comparatively high risk of osteoarthritis, a subject heterozygous for an allele associated with an increased risk of osteoarthritis is at a comparatively intermediate risk of osteoarthritis, and a subject homozygous for an allele associated with a decreased risk of osteoarthritis is at a comparatively low risk of osteoarthritis. A genotype may be assessed for a complementary strand, such that the complementary nucleotide at a particular position is detected.

[0094] Also featured are methods for determining risk of osteoarthritis and/or identifying a subject at risk of osteoarthritis by contacting a polypeptide or protein encoded by a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A from a subject with an antibody that specifically binds to an epitope associated with increased risk of osteoarthritis in the polypeptide.

Applications of Prognostic and Diagnostic Results to Pharmacogenomic Methods

[0095] Pharmacogenomics is a discipline that involves tailoring a treatment for a subject according to the subject's genotype as a particular treatment regimen may exert a differential effect depending upon the subject's genotype. For example, based upon the outcome of a prognostic test described herein, a clinician or physician may target pertinent information and preventative or therapeutic treatments to a subject who would be benefited by the information or treatment and avoid directing such information and treatments to a subject who would not be benefited (*e.g.*, the treatment has no therapeutic effect and/or the subject experiences adverse side effects).

[0096] The following is an example of a pharmacogenomic embodiment. A particular treatment regimen can exert a differential effect depending upon the subject's genotype. Where a candidate therapeutic exhibits a significant interaction with a major allele and a comparatively weak interaction with a minor allele (*e.g.*, an order of magnitude or greater difference in the interaction), such a therapeutic typically would not be administered to a subject genotyped as being homozygous for the minor allele, and sometimes not administered to a subject genotyped as being heterozygous for the minor allele. In another example, where a candidate therapeutic is not significantly toxic when administered to subjects who are homozygous for a major allele but is comparatively toxic when administered to subjects

heterozygous or homozygous for a minor allele, the candidate therapeutic is not typically administered to subjects who are genotyped as being heterozygous or homozygous with respect to the minor allele.

[0097] The methods described herein are applicable to pharmacogenomic methods for preventing, alleviating or treating osteoarthritis. For example, a nucleic acid sample from an individual may be subjected to a prognostic test described herein. Where one or more polymorphic variations associated with increased risk of osteoarthritis are identified in a subject, information for preventing or treating osteoarthritis and/or one or more osteoarthritis treatment regimens then may be prescribed to that subject.

[0098] In certain embodiments, a treatment or preventative regimen is specifically prescribed and/or administered to individuals who will most benefit from it based upon their risk of developing osteoarthritis assessed by the methods described herein. Thus, provided are methods for identifying a subject predisposed to osteoarthritis and then prescribing a therapeutic or preventative regimen to individuals identified as having a predisposition. Thus, certain embodiments are directed to a method for reducing osteoarthritis in a subject, which comprises: detecting the presence or absence of a polymorphic variant associated with osteoarthritis in a nucleotide sequence in a nucleic acid sample from a subject, where the nucleotide sequence comprises a polynucleotide sequence selected from the group consisting of: (a) a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; (b) a nucleotide sequence which encodes a polypeptide consisting of an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A, or a nucleotide sequence about 90% or more identical to a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; and (d) a fragment of a polynucleotide sequence of (a), (b), or (c); and prescribing or administering a treatment regimen to a subject from whom the sample originated where the presence of a polymorphic variation associated with osteoarthritis is detected in the nucleotide sequence. In these methods, predisposition results may be utilized in combination with other test results to diagnose osteoarthritis.

[0099] Certain preventative treatments often are prescribed to subjects having a predisposition to osteoarthritis and where the subject is diagnosed with osteoarthritis or is diagnosed as having symptoms indicative of an early stage of osteoarthritis. The treatment sometimes is preventative (*e.g.*, is prescribed or administered to reduce the probability that osteoarthritis arises or progresses), sometimes is therapeutic, and sometimes delays, alleviates or halts the progression of osteoarthritis. Any known preventative or therapeutic treatment for alleviating or preventing the occurrence of osteoarthritis is prescribed and/or administered. For example, the treatment often is directed to decreasing pain and improving joint movement. Examples of OA treatments include exercises to keep joints flexible and improve muscle strength. Different medications to control pain, including corticosteroids and nonsteroidal anti-inflammatory drugs (NSAIDs, *e.g.*, Voltaren); cyclooxygenase-2 (COX-2) inhibitors

(e.g., Celebrex, Vioxx, Mobic, and Bextra); monoclonal antibodies (e.g., Remicade); tumor necrosis factor inhibitors (e.g., Enbrel); or injections of glucocorticoids, hyaluronic acid or chondroitin sulfate into joints that are inflamed and not responsive to NSAIDS. Orally administered chondroitin sulfate also may be used as a therapeutic, as it may increase hyaluronic acid levels and viscosity of synovial fluid, and decrease collagenase levels in synovial fluid. Also, glucosamine can serve as an OA therapeutic as delivering it into joints may inhibit enzymes involved in cartilage degradation and enhance the production of hyaluronic acid. For mild pain without inflammation, acetaminophen may be used. Other treatments include: heat/cold therapy for temporary pain relief; joint protection to prevent strain or stress on painful joints; surgery to relieve chronic pain in damaged joints; and weight control to prevent extra stress on weight-bearing joints.

[0100] As therapeutic approaches for treating osteoarthritis continue to evolve and improve, the goal of treatments for osteoarthritis related disorders is to intervene even before clinical signs first manifest. Thus, genetic markers associated with susceptibility to osteoarthritis prove useful for early diagnosis, prevention and treatment of osteoarthritis.

[0101] As osteoarthritis preventative and treatment information can be specifically targeted to subjects in need thereof (e.g., those at risk of developing osteoarthritis or those in an early stage of osteoarthritis), provided herein is a method for preventing or reducing the risk of developing osteoarthritis in a subject, which comprises: (a) detecting the presence or absence of a polymorphic variation associated with osteoarthritis at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject; (b) identifying a subject with a predisposition to osteoarthritis, whereby the presence of the polymorphic variation is indicative of a predisposition to osteoarthritis in the subject; and (c) if such a predisposition is identified, providing the subject with information about methods or products to prevent or reduce osteoarthritis or to delay the onset of osteoarthritis. Also provided is a method of targeting information or advertising to a subpopulation of a human population based on the subpopulation being genetically predisposed to a disease or condition, which comprises: (a) detecting the presence or absence of a polymorphic variation associated with osteoarthritis at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject; (b) identifying the subpopulation of subjects in which the polymorphic variation is associated with osteoarthritis; and (c) providing information only to the subpopulation of subjects about a particular product which may be obtained and consumed or applied by the subject to help prevent or delay onset of the disease or condition.

[0102] Pharmacogenomics methods also may be used to analyze and predict a response to osteoarthritis treatment or a drug. For example, if pharmacogenomics analysis indicates a likelihood that an individual will respond positively to osteoarthritis treatment with a particular drug, the drug may be administered to the individual. Conversely, if the analysis indicates that an individual is likely to respond negatively to treatment with a particular drug, an alternative course of treatment may be prescribed. A

negative response may be defined as either the absence of an efficacious response or the presence of toxic side effects. The response to a therapeutic treatment can be predicted in a background study in which subjects in any of the following populations are genotyped: a population that responds favorably to a treatment regimen, a population that does not respond significantly to a treatment regimen, and a population that responds adversely to a treatment regimen (*e.g.*, exhibits one or more side effects). These populations are provided as examples and other populations and subpopulations may be analyzed. Based upon the results of these analyses, a subject is genotyped to predict whether he or she will respond favorably to a treatment regimen, not respond significantly to a treatment regimen, or respond adversely to a treatment regimen.

[0103] The tests described herein also are applicable to clinical drug trials. One or more polymorphic variants indicative of response to an agent for treating osteoarthritis or to side effects to an agent for treating osteoarthritis may be identified using the methods described herein. Thereafter, potential participants in clinical trials of such an agent may be screened to identify those individuals most likely to respond favorably to the drug and exclude those likely to experience side effects. In that way, the effectiveness of drug treatment may be measured in individuals who respond positively to the drug, without lowering the measurement as a result of the inclusion of individuals who are unlikely to respond positively in the study and without risking undesirable safety problems.

[0104] Thus, another embodiment is a method of selecting an individual for inclusion in a clinical trial of a treatment or drug comprising the steps of: (a) obtaining a nucleic acid sample from an individual; (b) determining the identity of a polymorphic variation which is associated with a positive response to the treatment or the drug, or at least one polymorphic variation which is associated with a negative response to the treatment or the drug in the nucleic acid sample, and (c) including the individual in the clinical trial if the nucleic acid sample contains said polymorphic variation associated with a positive response to the treatment or the drug or if the nucleic acid sample lacks said polymorphic variation associated with a negative response to the treatment or the drug. In addition, the methods described herein for selecting an individual for inclusion in a clinical trial of a treatment or drug encompass methods with any further limitation described in this disclosure, or those following, specified alone or in any combination. The polymorphic variation may be in a sequence selected individually or in any combination from the group consisting of (i) a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; (ii) a nucleotide sequence which encodes a polypeptide consisting of an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; (iii) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A, or a nucleotide sequence about 90% or more identical to a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A; and (iv) a fragment of a polynucleotide sequence of (i), (ii), or (iii) comprising the polymorphic site. The including

step (c) optionally comprises administering the drug or the treatment to the individual if the nucleic acid sample contains the polymorphic variation associated with a positive response to the treatment or the drug and the nucleic acid sample lacks said biallelic marker associated with a negative response to the treatment or the drug.

[0105] Also provided herein is a method of partnering between a diagnostic/prognostic testing provider and a provider of a consumable product, which comprises: (a) the diagnostic/prognostic testing provider detects the presence or absence of a polymorphic variation associated with osteoarthritis at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject; (b) the diagnostic/prognostic testing provider identifies the subpopulation of subjects in which the polymorphic variation is associated with osteoarthritis; (c) the diagnostic/prognostic testing provider forwards information to the subpopulation of subjects about a particular product which may be obtained and consumed or applied by the subject to help prevent or delay onset of the disease or condition; and (d) the provider of a consumable product forwards to the diagnostic test provider a fee every time the diagnostic/prognostic test provider forwards information to the subject as set forth in step (c) above.

Compositions Comprising Osteoarthritis-Directed Molecules

[0106] Featured herein is a composition comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and one or more molecules specifically directed and targeted to a nucleic acid comprising a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence, other nucleotide sequence referenced in Table A, or an encoded amino acid sequence. Such directed molecules include, but are not limited to, a compound that binds to a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence, or other nucleotide sequence referenced in Table A, or encoded amino acid sequence; a RNAi or siRNA molecule having a strand complementary or substantially complementary to a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A (e.g., hybridizes to a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A under conditions of high stringency); an antisense nucleic acid complementary or substantially complementary to an RNA encoded by a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A (e.g., hybridizes to a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A under conditions of high stringency); a ribozyme that hybridizes to a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A (e.g., hybridizes to a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A under conditions of high stringency); a nucleic acid aptamer that specifically binds a polypeptide encoded by a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A; and an antibody that specifically binds to a polypeptide encoded by a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A or

binds to a nucleic acid having such a nucleotide sequence. In an embodiment, the antibody selectively binds to an epitope comprising an amino acid encoded by rs734784, rs1042164, rs749670, rs955592 and rs1040461. In specific embodiments, the osteoarthritis directed molecule interacts with a nucleic acid or polypeptide variant associated with osteoarthritis, such as variants referenced herein. In other embodiments, the osteoarthritis directed molecule interacts with a polypeptide involved in a signal pathway of a polypeptide encoded by a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A, or a nucleic acid comprising such a nucleotide sequence.

[0107] Compositions sometimes include an adjuvant known to stimulate an immune response, and in certain embodiments, an adjuvant that stimulates a T-cell lymphocyte response. Adjuvants are known, including but not limited to an aluminum adjuvant (e.g., aluminum hydroxide); a cytokine adjuvant or adjuvant that stimulates a cytokine response (e.g., interleukin (IL)-12 and/or gamma-interferon cytokines); a Freund-type mineral oil adjuvant emulsion (e.g., Freund's complete or incomplete adjuvant); a synthetic lipid compound; a copolymer adjuvant (e.g., TitreMax); a saponin; Quil A; a liposome; an oil-in-water emulsion (e.g., an emulsion stabilized by Tween 80 and pluronic polyoxyethylene/polyoxypropylene block copolymer (Syntex Adjuvant Formulation); TitreMax; detoxified endotoxin (MPL) and mycobacterial cell wall components (TDW, CWS) in 2% squalene (Ribi Adjuvant System)); a muramyl dipeptide; an immune-stimulating complex (ISCOM, e.g., an Ag-modified saponin/cholesterol micelle that forms stable cage-like structure); an aqueous phase adjuvant that does not have a depot effect (e.g., Gerbu adjuvant); a carbohydrate polymer (e.g., AdjuPrime); L-tyrosine; a manide-oleate compound (e.g., Montanide); an ethylene-vinyl acetate copolymer (e.g., Elvax 40W1,2); or lipid A, for example. Such compositions are useful for generating an immune response against osteoarthritis directed molecule (e.g., an HLA-binding subsequence within a polypeptide encoded by a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence). In such methods, a peptide having an amino acid subsequence of a polypeptide encoded by a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence is delivered to a subject, where the subsequence binds to an HLA molecule and induces a CTL lymphocyte response. The peptide sometimes is delivered to the subject as an isolated peptide or as a minigene in a plasmid that encodes the peptide. Methods for identifying HLA-binding subsequences in such polypeptides are known (see e.g., publication WO02/20616 and PCT application US98/01373 for methods of identifying such sequences).

[0108] The cell may be in a group of cells cultured *in vitro* or in a tissue maintained *in vitro* or present in an animal *in vivo* (e.g., a rat, mouse, ape or human). In certain embodiments, a composition comprises a component from a cell such as a nucleic acid molecule (e.g., genomic DNA), a protein mixture or isolated protein, for example. The aforementioned compositions have utility in diagnostic, prognostic and pharmacogenomic methods described previously and in therapeutics described hereafter. Certain osteoarthritis directed molecules are described in greater detail below.

Compounds

[0109] Compounds can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; peptoid libraries (libraries of molecules having the functionalities of peptides, but with a novel, non-peptide backbone which are resistant to enzymatic degradation but which nevertheless remain bioactive (see, e.g., Zuckermann et al., J. Med. Chem. 37: 2678-85 (1994)); spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; "one-bead one-compound" library methods; and synthetic library methods using affinity chromatography selection. Biological library and peptoid library approaches are typically limited to peptide libraries, while the other approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, Anticancer Drug Des. 12: 145, (1997)). Examples of methods for synthesizing molecular libraries are described, for example, in DeWitt et al., Proc. Natl. Acad. Sci. U.S.A. 90: 6909 (1993); Erb et al., Proc. Natl. Acad. Sci. USA 91: 11422 (1994); Zuckermann et al., J. Med. Chem. 37: 2678 (1994); Cho et al., Science 261: 1303 (1993); Carrell et al., Angew. Chem. Int. Ed. Engl. 33: 2059 (1994); Carell et al., Angew. Chem. Int. Ed. Engl. 33: 2061 (1994); and in Gallop et al., J. Med. Chem. 37: 1233 (1994).

[0110] Libraries of compounds may be presented in solution (e.g., Houghten, Biotechniques 13: 412-421 (1992)), or on beads (Lam, Nature 354: 82-84 (1991)), chips (Fodor, Nature 364: 555-556 (1993)), bacteria or spores (Ladner, United States Patent No. 5,223,409), plasmids (Cull et al., Proc. Natl. Acad. Sci. USA 89: 1865-1869 (1992)) or on phage (Scott and Smith, Science 249: 386-390 (1990); Devlin, Science 249: 404-406 (1990); Cwirla et al., Proc. Natl. Acad. Sci. 87: 6378-6382 (1990); Felici, J. Mol. Biol. 222: 301-310 (1991); Ladner supra.).

[0111] A compound sometimes alters expression and sometimes alters activity of a polypeptide target and may be a small molecule. Small molecules include, but are not limited to, peptides, peptidomimetics (e.g., peptoids), amino acids, amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs, organic or inorganic compounds (i.e., including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 1,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds.

Antisense Nucleic Acid Molecules, Ribozymes, RNAi, siRNA and Modified Nucleic Acid Molecules

[0112] An "antisense" nucleic acid refers to a nucleotide sequence complementary to a "sense" nucleic acid encoding a polypeptide, e.g., complementary to the coding strand of a double-stranded

cDNA molecule or complementary to an mRNA sequence. The antisense nucleic acid can be complementary to an entire coding strand, or to a portion thereof or a substantially identical sequence thereof. In another embodiment, the antisense nucleic acid molecule is antisense to a “noncoding region” of the coding strand of a nucleotide sequence (e.g., 5’ and 3’ untranslated regions in SEQ ID NO: 1-3 or a nucleotide sequence referenced in Table A).

[0113] An antisense nucleic acid can be designed such that it is complementary to the entire coding region of an mRNA encoded by a nucleotide sequence (e.g., SEQ ID NO: 1-3 or a nucleotide sequence referenced in Table A), and often the antisense nucleic acid is an oligonucleotide antisense to only a portion of a coding or noncoding region of the mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of the mRNA, e.g., between the -10 and +10 regions of the target gene nucleotide sequence of interest. An antisense oligonucleotide can be, for example, about 7, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, or more nucleotides in length. The antisense nucleic acids, which include the ribozymes described hereafter, can be designed to target a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence, often a variant associated with osteoarthritis, or a substantially identical sequence thereof. Among the variants, minor alleles and major alleles can be targeted, and those associated with a higher risk of osteoarthritis are often designed, tested, and administered to subjects.

[0114] An antisense nucleic acid can be constructed using chemical synthesis and enzymatic ligation reactions using standard procedures. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used. Antisense nucleic acid also can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

[0115] When utilized as therapeutics, antisense nucleic acids typically are administered to a subject (e.g., by direct injection at a tissue site) or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a polypeptide and thereby inhibit expression of the polypeptide, for example, by inhibiting transcription and/or translation. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then are administered systemically. For systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, for example, by linking antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. Antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. Sufficient intracellular

concentrations of antisense molecules are achieved by incorporating a strong promoter, such as a pol II or pol III promoter, in the vector construct.

[0116] Antisense nucleic acid molecules sometimes are alpha-anomeric nucleic acid molecules. An alpha-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual beta-units, the strands run parallel to each other (Gaultier et al., *Nucleic Acids. Res.* 15: 6625-6641 (1987)). Antisense nucleic acid molecules can also comprise a 2'-o-methylribonucleotide (Inoue et al., *Nucleic Acids Res.* 15: 6131-6148 (1987)) or a chimeric RNA-DNA analogue (Inoue et al., *FEBS Lett.* 215: 327-330 (1987)). Antisense nucleic acids sometimes are composed of DNA or PNA or any other nucleic acid derivatives described previously.

[0117] In another embodiment, an antisense nucleic acid is a ribozyme. A ribozyme having specificity for a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A can include one or more sequences complementary to such a nucleotide sequence, and a sequence having a known catalytic region responsible for mRNA cleavage (see e.g., U.S. Pat. No. 5,093,246 or Haselhoff and Gerlach, *Nature* 334: 585-591 (1988)). For example, a derivative of a *Tetrahymena* L-19 IVS RNA is sometimes utilized in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a mRNA (see e.g., Cech et al. U.S. Patent No. 4,987,071; and Cech et al. U.S. Patent No. 5,116,742). Also, target mRNA sequences can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules (see e.g., Bartel & Szostak, *Science* 261: 1411-1418 (1993)).

[0118] Osteoarthritis directed molecules include in certain embodiments nucleic acids that can form triple helix structures with a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A, or a substantially identical sequence thereof, especially one that includes a regulatory region that controls expression of a polypeptide. Gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of a nucleotide sequence referenced herein or a substantially identical sequence (e.g., promoter and/or enhancers) to form triple helical structures that prevent transcription of a gene in target cells (see e.g., Helene, *Anticancer Drug Des.* 6(6): 569-84 (1991); Helene et al., *Ann. N.Y. Acad. Sci.* 660: 27-36 (1992); and Maher, *Bioassays* 14(12): 807-15 (1992). Potential sequences that can be targeted for triple helix formation can be increased by creating a so-called "switchback" nucleic acid molecule. Switchback molecules are synthesized in an alternating 5'-3', 3'-5' manner, such that they base pair with first one strand of a duplex and then the other, eliminating the necessity for a sizeable stretch of either purines or pyrimidines to be present on one strand of a duplex.

[0119] Osteoarthritis directed molecules include RNAi and siRNA nucleic acids. Gene expression may be inhibited by the introduction of double-stranded RNA (dsRNA), which induces potent and specific gene silencing, a phenomenon called RNA interference or RNAi. See, e.g., Fire et al., US Patent

Number 6,506,559; Tuschl et al. PCT International Publication No. WO 01/75164; Kay et al. PCT International Publication No. WO 03/010180A1; or Bosher JM, Labouesse, Nat Cell Biol 2000 Feb;2(2):E31-6. This process has been improved by decreasing the size of the double-stranded RNA to 20-24 base pairs (to create small-interfering RNAs or siRNAs) that “switched off” genes in mammalian cells without initiating an acute phase response, i.e., a host defense mechanism that often results in cell death (see, e.g., Caplen et al. Proc Natl Acad Sci U S A. 2001 Aug 14;98(17):9742-7 and Elbashir et al. Methods 2002 Feb;26(2):199-213). There is increasing evidence of post-transcriptional gene silencing by RNA interference (RNAi) for inhibiting targeted expression in mammalian cells at the mRNA level, in human cells. There is additional evidence of effective methods for inhibiting the proliferation and migration of tumor cells in human patients, and for inhibiting metastatic cancer development (see, e.g., U.S. Patent Application No. US2001000993183; Caplen et al. Proc Natl Acad Sci U S A; and Abderrahmani et al. Mol Cell Biol 2001 Nov21(21):7256-67).

[0120] An “siRNA” or “RNAi” refers to a nucleic acid that forms a double stranded RNA and has the ability to reduce or inhibit expression of a gene or target gene when the siRNA is delivered to or expressed in the same cell as the gene or target gene. “siRNA” refers to short double-stranded RNA formed by the complementary strands. Complementary portions of the siRNA that hybridize to form the double stranded molecule often have substantial or complete identity to the target molecule sequence. In one embodiment, an siRNA refers to a nucleic acid that has substantial or complete identity to a target gene and forms a double stranded siRNA.

[0121] When designing the siRNA molecules, the targeted region often is selected from a given DNA sequence beginning 50 to 100 nucleotides downstream of the start codon. See, e.g., Elbashir et al., Methods 26:199-213 (2002). Initially, 5’ or 3’ UTRs and regions nearby the start codon were avoided assuming that UTR-binding proteins and/or translation initiation complexes may interfere with binding of the siRNP or RISC endonuclease complex. Sometimes regions of the target 23 nucleotides in length conforming to the sequence motif AA(N19)TT (N, an nucleotide), and regions with approximately 30% to 70% G/C-content (often about 50% G/C-content) often are selected. If no suitable sequences are found, the search often is extended using the motif NA(N21). The sequence of the sense siRNA sometimes corresponds to (N19) TT or N21 (position 3 to 23 of the 23-nt motif), respectively. In the latter case, the 3’ end of the sense siRNA often is converted to TT. The rationale for this sequence conversion is to generate a symmetric duplex with respect to the sequence composition of the sense and antisense 3’ overhangs. The antisense siRNA is synthesized as the complement to position 1 to 21 of the 23-nt motif. Because position 1 of the 23-nt motif is not recognized sequence-specifically by the antisense siRNA, the 3’-most nucleotide residue of the antisense siRNA can be chosen deliberately. However, the penultimate nucleotide of the antisense siRNA (complementary to position 2 of the 23-nt motif) often is complementary to the targeted sequence. For simplifying chemical synthesis, TT often is

utilized. siRNAs corresponding to the target motif NAR(N17)YNN, where R is purine (A,G) and Y is pyrimidine (C,U), often are selected. Respective 21 nucleotide sense and antisense siRNAs often begin with a purine nucleotide and can also be expressed from pol III expression vectors without a change in targeting site. Expression of RNAs from pol III promoters often is efficient when the first transcribed nucleotide is a purine.

[0122] The sequence of the siRNA can correspond to the full length target gene, or a subsequence thereof. Often, the siRNA is about 15 to about 50 nucleotides in length (e.g., each complementary sequence of the double stranded siRNA is 15-50 nucleotides in length, and the double stranded siRNA is about 15-50 base pairs in length, sometimes about 20-30 nucleotides in length or about 20-25 nucleotides in length, e.g., 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides in length. The siRNA sometimes is about 21 nucleotides in length. Methods of using siRNA are well known in the art, and specific siRNA molecules may be purchased from a number of companies including Dharmacon Research, Inc.

[0123] Antisense, ribozyme, RNAi and siRNA nucleic acids can be altered to form modified nucleic acid molecules. The nucleic acids can be altered at base moieties, sugar moieties or phosphate backbone moieties to improve stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of nucleic acid molecules can be modified to generate peptide nucleic acids (see Hyrup et al., *Bioorganic & Medicinal Chemistry* 4 (1): 5-23 (1996)). As used herein, the terms “peptide nucleic acid” or “PNA” refers to a nucleic acid mimic such as a DNA mimic, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of a PNA can allow for specific hybridization to DNA and RNA under conditions of low ionic strength. Synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described, for example, in Hyrup et al., (1996) *supra* and Perry-O’Keefe et al., *Proc. Natl. Acad. Sci.* 93: 14670-675 (1996).

[0124] PNA nucleic acids can be used in prognostic, diagnostic, and therapeutic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation arrest or inhibiting replication. PNA nucleic acid molecules can also be used in the analysis of single base pair mutations in a gene, (e.g., by PNA-directed PCR clamping); as “artificial restriction enzymes” when used in combination with other enzymes, (e.g., S1 nucleases (Hyrup (1996) *supra*)); or as probes or primers for DNA sequencing or hybridization (Hyrup et al., (1996) *supra*; Perry-O’Keefe *supra*).

[0125] In other embodiments, oligonucleotides may include other appended groups such as peptides (e.g., for targeting host cell receptors *in vivo*), or agents facilitating transport across cell membranes (see e.g., Letsinger et al., *Proc. Natl. Acad. Sci. USA* 86: 6553-6556 (1989); Lemaitre et al., *Proc. Natl. Acad. Sci. USA* 84: 648-652 (1987); PCT Publication No. W088/09810) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization-

triggered cleavage agents (See, e.g., Krol et al., Bio-Techniques 6: 958-976 (1988)) or intercalating agents. (See, e.g., Zon, Pharm. Res. 5: 539-549 (1988)). To this end, the oligonucleotide may be conjugated to another molecule, (e.g., a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent).

[0126] Also included herein are molecular beacon oligonucleotide primer and probe molecules having one or more regions complementary to a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A, or a substantially identical sequence thereof, two complementary regions one having a fluorophore and one a quencher such that the molecular beacon is useful for quantifying the presence of the nucleic acid in a sample. Molecular beacon nucleic acids are described, for example, in Lizardi et al., U.S. Patent No. 5,854,033; Nazarenko et al., U.S. Patent No. 5,866,336, and Livak et al., U.S. Patent 5,876,930.

Antibodies

[0127] The term “antibody” as used herein refers to an immunoglobulin molecule or immunologically active portion thereof, i.e., an antigen-binding portion. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and F(ab')₂ fragments which can be generated by treating the antibody with an enzyme such as pepsin. An antibody sometimes is a polyclonal, monoclonal, recombinant (e.g., a chimeric or humanized), fully human, non-human (e.g., murine), or a single chain antibody. An antibody may have effector function and can fix complement, and is sometimes coupled to a toxin or imaging agent.

[0128] A full-length polypeptide or antigenic peptide fragment encoded by a nucleotide sequence referenced herein can be used as an immunogen or can be used to identify antibodies made with other immunogens, e.g., cells, membrane preparations, and the like. An antigenic peptide often includes at least 8 amino acid residues of the amino acid sequences encoded by a nucleotide sequence referenced herein, or substantially identical sequence thereof, and encompasses an epitope. Antigenic peptides sometimes include 10 or more amino acids, 15 or more amino acids, 20 or more amino acids, or 30 or more amino acids. Hydrophilic and hydrophobic fragments of polypeptides sometimes are used as immunogens.

[0129] Epitopes encompassed by the antigenic peptide are regions located on the surface of the polypeptide (e.g., hydrophilic regions) as well as regions with high antigenicity. For example, an Emini surface probability analysis of the human polypeptide sequence can be used to indicate the regions that have a particularly high probability of being localized to the surface of the polypeptide and are thus likely to constitute surface residues useful for targeting antibody production. The antibody may bind an epitope on any domain or region on polypeptides described herein.

[0130] Also, chimeric, humanized, and completely human antibodies are useful for applications which include repeated administration to subjects. Chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, can be made using standard recombinant DNA techniques. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in Robinson et al International Application No. PCT/US86/02269; Akira, et al European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison et al European Patent Application 173,494; Neuberger et al PCT International Publication No. WO 86/01533; Cabilly et al U.S. Patent No. 4,816,567; Cabilly et al European Patent Application 125,023; Better et al., Science 240: 1041-1043 (1988); Liu et al., Proc. Natl. Acad. Sci. USA 84: 3439-3443 (1987); Liu et al., J. Immunol. 139: 3521-3526 (1987); Sun et al., Proc. Natl. Acad. Sci. USA 84: 214-218 (1987); Nishimura et al., Canc. Res. 47: 999-1005 (1987); Wood et al., Nature 314: 446-449 (1985); and Shaw et al., J. Natl. Cancer Inst. 80: 1553-1559 (1988); Morrison, S. L., Science 229: 1202-1207 (1985); Oi et al., BioTechniques 4: 214 (1986); Winter U.S. Patent 5,225,539; Jones et al., Nature 321: 552-525 (1986); Verhoeyan et al., Science 239: 1534; and Beidler et al., J. Immunol. 141: 4053-4060 (1988).

[0131] Completely human antibodies are particularly desirable for therapeutic treatment of human patients. Such antibodies can be produced using transgenic mice that are incapable of expressing endogenous immunoglobulin heavy and light chains genes, but which can express human heavy and light chain genes. See, for example, Lonberg and Huszar, Int. Rev. Immunol. 13: 65-93 (1995); and U.S. Patent Nos. 5,625,126; 5,633,425; 5,569,825; 5,661,016; and 5,545,806. In addition, companies such as Abgenix, Inc. (Fremont, CA) and Medarex, Inc. (Princeton, NJ), can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above. Completely human antibodies that recognize a selected epitope also can be generated using a technique referred to as "guided selection." In this approach a selected non-human monoclonal antibody (e.g., a murine antibody) is used to guide the selection of a completely human antibody recognizing the same epitope. This technology is described for example by Jespers et al., Bio/Technology 12: 899-903 (1994).

[0132] An antibody can be a single chain antibody. A single chain antibody (scFV) can be engineered (see, e.g., Colcher et al., Ann. N Y Acad. Sci. 880: 263-80 (1999); and Reiter, Clin. Cancer Res. 2: 245-52 (1996)). Single chain antibodies can be dimerized or multimerized to generate multivalent antibodies having specificities for different epitopes of the same target polypeptide.

[0133] Antibodies also may be selected or modified so that they exhibit reduced or no ability to bind an Fc receptor. For example, an antibody may be an isotype or subtype, fragment or other mutant, which does not support binding to an Fc receptor (e.g., it has a mutagenized or deleted Fc receptor binding region).

[0134] Also, an antibody (or fragment thereof) may be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1 dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (e.g., methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (e.g., mechlorethamine, thiotepa chlorambucil, melphalan, carmustine (BCNU) and lomustine (CCNU), cyclophosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (e.g., daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (e.g., dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (e.g., vincristine and vinblastine).

[0135] Antibody conjugates can be used for modifying a given biological response. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a polypeptide such as tumor necrosis factor, gamma-interferon, alpha-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophage colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors. Also, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980, for example.

[0136] An antibody (e.g., monoclonal antibody) can be used to isolate target polypeptides by standard techniques, such as affinity chromatography or immunoprecipitation. Moreover, an antibody can be used to detect a target polypeptide (e.g., in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the polypeptide. Antibodies can be used diagnostically to monitor polypeptide levels in tissue as part of a clinical testing procedure, e.g., to determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (i.e., physically linking) the antibody to a detectable substance (i.e., antibody labeling). Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol;

examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I , ^{131}I , ^{35}S or ^3H . Also, an antibody can be utilized as a test molecule for determining whether it can treat osteoarthritis, and as a therapeutic for administration to a subject for treating osteoarthritis.

[0137] An antibody can be made by immunizing with a purified antigen, or a fragment thereof, e.g., a fragment described herein, a membrane associated antigen, tissues, e.g., crude tissue preparations, whole cells, preferably living cells, lysed cells, or cell fractions.

[0138] Included herein are antibodies which bind only a native polypeptide, only denatured or otherwise non-native polypeptide, or which bind both, as well as those having linear or conformational epitopes. Conformational epitopes sometimes can be identified by selecting antibodies that bind to native but not denatured polypeptide. Also featured are antibodies that specifically bind to a polypeptide variant associated with osteoarthritis.

Methods for Identifying Candidate Therapeutics for Treating Osteoarthritis

[0139] Current therapies for the treatment of osteoarthritis have limited efficacy, limited tolerability and significant mechanism-based side effects, and few of the available therapies adequately address underlying defects. Current therapeutic approaches were largely developed in the absence of defined molecular targets or even a solid understanding of disease pathogenesis. Therefore, provided are methods of identifying candidate therapeutics that target biochemical pathways related to the development of osteoarthritis.

[0140] Thus, featured herein are methods for identifying a candidate therapeutic for treating osteoarthritis. The methods comprise contacting a test molecule with a target molecule in a system. A “target molecule” as used herein refers to a *KIAA0296*, *chrom 4* or *SNW1* nucleic acid or other nucleotide sequence referenced in Table A, a substantially identical nucleic acid thereof, or a fragment thereof, and an encoded polypeptide of the foregoing. The methods also comprise determining the presence or absence of an interaction between the test molecule and the target molecule, where the presence of an interaction between the test molecule and the nucleic acid or polypeptide identifies the test molecule as a candidate osteoarthritis therapeutic. The interaction between the test molecule and the target molecule may be quantified.

[0141] Test molecules and candidate therapeutics include, but are not limited to, compounds, antisense nucleic acids, siRNA molecules, ribozymes, polypeptides or proteins encoded by a *KIAA0296*, *chrom 4* or *SNW1* nucleotide sequence or other nucleotide sequence referenced in Table A, or a substantially identical sequence or fragment thereof, and immunotherapeutics (e.g., antibodies and HLA-presented polypeptide fragments). A test molecule or candidate therapeutic may act as a modulator of target molecule concentration or target molecule function in a system. A “modulator” may agonize (i.e.,

up-regulates) or antagonize (i.e., down-regulates) a target molecule concentration partially or completely in a system by affecting such cellular functions as DNA replication and/or DNA processing (e.g., DNA methylation or DNA repair), RNA transcription and/or RNA processing (e.g., removal of intronic sequences and/or translocation of spliced mRNA from the nucleus), polypeptide production (e.g., translation of the polypeptide from mRNA), and/or polypeptide post-translational modification (e.g., glycosylation, phosphorylation, and proteolysis of pro-polypeptides). A modulator may also agonize or antagonize a biological function of a target molecule partially or completely, where the function may include adopting a certain structural conformation, interacting with one or more binding partners, ligand binding, catalysis (e.g., phosphorylation, dephosphorylation, hydrolysis, methylation, and isomerization), and an effect upon a cellular event (e.g., effecting progression of osteoarthritis).

[0142] As used herein, the term “system” refers to a cell free *in vitro* environment and a cell-based environment such as a collection of cells, a tissue, an organ, or an organism. A system is “contacted” with a test molecule in a variety of manners, including adding molecules in solution and allowing them to interact with one another by diffusion, cell injection, and any administration routes in an animal. As used herein, the term “interaction” refers to an effect of a test molecule on test molecule, where the effect sometimes is binding between the test molecule and the target molecule, and sometimes is an observable change in cells, tissue, or organism.

[0143] There are many standard methods for detecting the presence or absence of interaction between a test molecule and a target molecule. For example, titrametric, acidimetric, radiometric, NMR, monolayer, polarographic, spectrophotometric, fluorescent, and ESR assays probative of a target molecule interaction may be utilized. Any modulator can be tested in such methods and modulators for certain targets described in Table A are known. For example, modulators of protein tyrosine phosphatases (e.g., *PTPNI* includes a protein phosphatase domain) are described in WO-03072537, WO-03020688, WO-00218321, WO-00218323, WO-03055883, WO-03041729, WO-00226707, WO-00226743 and WO-03037328; modulators of potassium channels (e.g., *KCNSI* includes a potassium channel domain) are described in WO-09962891, WO-09716437, WO-09521813, WO-09521823, WO-09521824, WO-09521825 and WO-03088908; modulators of annexin (e.g., *ANXA6* includes an annexin domain) are described in WO-2004018670, WO-02067857, WO-2004013303 and WO-00147510; and modulators of protein kinases (e.g., *FYN* is a protein kinase) are described in WO-03081210, WO-02080926, WO-02076986, WO-03077921, WO03026666, WO03026665 and WO03026664.

[0144] Test molecule/target molecule interactions can be detected and/or quantified using assays known in the art. For example, an interaction can be determined by labeling the test molecule and/or the target molecule, where the label is covalently or non-covalently attached to the test molecule or target molecule. The label is sometimes a radioactive molecule such as ^{125}I , ^{131}I , ^{35}S or ^3H , which can be detected by direct counting of radioemission or by scintillation counting. Also, enzymatic labels such as

horseradish peroxidase, alkaline phosphatase, or luciferase may be utilized where the enzymatic label can be detected by determining conversion of an appropriate substrate to product. In addition, presence or absence of an interaction can be determined without labeling. For example, a microphysiometer (*e.g.*, Cytosensor) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indication of an interaction between a test molecule and target molecule (McConnell, H. M. *et al.*, *Science* 257: 1906-1912 (1992)).

[0145] In cell-based systems, cells typically include a *KIAA0296*, *chrom 4* or *SNW1* nucleic acid or other nucleotide sequence referenced in Table A, an encoded polypeptide, or substantially identical nucleic acid or polypeptide thereof, and are often of mammalian origin, although the cell can be of any origin. Whole cells, cell homogenates, and cell fractions (*e.g.*, cell membrane fractions) can be subjected to analysis. Where interactions between a test molecule with a target polypeptide are monitored, soluble and/or membrane bound forms of the polypeptide may be utilized. Where membrane-bound forms of the polypeptide are used, it may be desirable to utilize a solubilizing agent. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®, Isotridecypoly(ethylene glycol ether)_n, 3-[(3-cholamidopropyl)dimethylamminio]-1-propane sulfonate (CHAPS), 3-[(3-cholamidopropyl)dimethylamminio]-2-hydroxy-1-propane sulfonate (CHAPSO), or N-dodecyl-N,N-dimethyl-3-ammonio-1-propane sulfonate.

[0146] An interaction between a test molecule and target molecule also can be detected by monitoring fluorescence energy transfer (FET) (*see, e.g.*, Lakowicz *et al.*, U.S. Patent No. 5,631,169; Stavrianopoulos *et al.* U.S. Patent No. 4,868,103). A fluorophore label on a first, "donor" molecule is selected such that its emitted fluorescent energy will be absorbed by a fluorescent label on a second, "acceptor" molecule, which in turn is able to fluoresce due to the absorbed energy. Alternately, the "donor" polypeptide molecule may simply utilize the natural fluorescent energy of tryptophan residues. Labels are chosen that emit different wavelengths of light, such that the "acceptor" molecule label may be differentiated from that of the "donor". Since the efficiency of energy transfer between the labels is related to the distance separating the molecules, the spatial relationship between the molecules can be assessed. In a situation in which binding occurs between the molecules, the fluorescent emission of the "acceptor" molecule label in the assay should be maximal. An FET binding event can be conveniently measured through standard fluorometric detection means well known in the art (*e.g.*, using a fluorimeter).

[0147] In another embodiment, determining the presence or absence of an interaction between a test molecule and a target molecule can be effected by monitoring surface plasmon resonance (*see, e.g.*, Sjolander & Urbanicz, *Anal. Chem.* 63: 2338-2345 (1991) and Szabo *et al.*, *Curr. Opin. Struct. Biol.* 5:

699-705 (1995)). “Surface plasmon resonance” or “biomolecular interaction analysis (BIA)” can be utilized to detect biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore). Changes in the mass at the binding surface (indicative of a binding event) result in alterations of the refractive index of light near the surface (the optical phenomenon of surface plasmon resonance (SPR)), resulting in a detectable signal which can be used as an indication of real-time reactions between biological molecules.

[0148] In another embodiment, the target molecule or test molecules are anchored to a solid phase, facilitating the detection of target molecule/test molecule complexes and separation of the complexes from free, uncomplexed molecules. The target molecule or test molecule is immobilized to the solid support. In an embodiment, the target molecule is anchored to a solid surface, and the test molecule, which is not anchored, can be labeled, either directly or indirectly, with detectable labels discussed herein.

[0149] It may be desirable to immobilize a target molecule, an anti-target molecule antibody, and/or test molecules to facilitate separation of target molecule/test molecule complexes from uncomplexed forms, as well as to accommodate automation of the assay. The attachment between a test molecule and/or target molecule and the solid support may be covalent or non-covalent (*see, e.g.*, U.S. Patent No. 6,022,688 for non-covalent attachments). The solid support may be one or more surfaces of the system, such as one or more surfaces in each well of a microtiter plate, a surface of a silicon wafer, a surface of a bead (*see, e.g.*, Lam, *Nature* 354: 82-84 (1991)) that is optionally linked to another solid support, or a channel in a microfluidic device, for example. Types of solid supports, linker molecules for covalent and non-covalent attachments to solid supports, and methods for immobilizing nucleic acids and other molecules to solid supports are well known (*see, e.g.*, U.S. Patent Nos. 6,261,776; 5,900,481; 6,133,436; and 6,022,688; and WIPO publication WO 01/18234).

[0150] In an embodiment, target molecule may be immobilized to surfaces via biotin and streptavidin. For example, biotinylated target polypeptide can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). In another embodiment, a target polypeptide can be prepared as a fusion polypeptide. For example, glutathione-S-transferase/target polypeptide fusion can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivitized microtiter plates, which are then combined with a test molecule under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, or the matrix is immobilized in the case of beads, and complex formation is determined directly or indirectly as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of target molecule binding or activity is determined using standard techniques.

[0151] In an embodiment, the non-immobilized component is added to the coated surface containing the anchored component. After the reaction is complete, unreacted components are removed (*e.g.*, by washing) under conditions such that a significant percentage of complexes formed will remain immobilized to the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of manners. Where the previously non-immobilized component is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the previously non-immobilized component is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface, *e.g.*, by adding a labeled antibody specific for the immobilized component, where the antibody, in turn, can be directly labeled or indirectly labeled with, *e.g.*, a labeled anti-Ig antibody.

[0152] In another embodiment, an assay is performed utilizing antibodies that specifically bind target molecule or test molecule but do not interfere with binding of the target molecule to the test molecule. Such antibodies can be derivitized to a solid support, and unbound target molecule may be immobilized by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the target molecule.

[0153] Cell free assays also can be conducted in a liquid phase. In such an assay, reaction products are separated from unreacted components, by any of a number of standard techniques, including but not limited to: differential centrifugation (*see, e.g.*, Rivas, G., and Minton, *Trends Biochem Sci Aug; 18(8): 284-7 (1993)*); chromatography (gel filtration chromatography, ion-exchange chromatography); electrophoresis (*see, e.g.*, Ausubel *et al.*, eds. *Current Protocols in Molecular Biology*, J. Wiley: New York (1999)); and immunoprecipitation (*see, e.g.*, Ausubel *et al.*, eds., *supra*). Media and chromatographic techniques are known to one skilled in the art (*see, e.g.*, Heegaard, *J Mol. Recognit. Winter; 11(1-6): 141-8 (1998)*; Hage & Tweed, *J. Chromatogr. B Biomed. Sci. Appl. Oct 10; 699 (1-2): 499-525 (1997)*). Further, fluorescence energy transfer may also be conveniently utilized, as described herein, to detect binding without further purification of the complex from solution.

[0154] In another embodiment, modulators of target molecule expression are identified. For example, a cell or cell free mixture is contacted with a candidate compound and the expression of target mRNA or target polypeptide is evaluated relative to the level of expression of target mRNA or target polypeptide in the absence of the candidate compound. When expression of target mRNA or target polypeptide is greater in the presence of the candidate compound than in its absence, the candidate compound is identified as an agonist of target mRNA or target polypeptide expression. Alternatively, when expression of target mRNA or target polypeptide is less (*e.g.*, less with statistical significance) in the presence of the candidate compound than in its absence, the candidate compound is identified as an

antagonist or inhibitor of target mRNA or target polypeptide expression. The level of target mRNA or target polypeptide expression can be determined by methods described herein.

[0155] In another embodiment, binding partners that interact with a target molecule are detected. The target molecules can interact with one or more cellular or extracellular macromolecules, such as polypeptides *in vivo*, and these interacting molecules are referred to herein as “binding partners.” Binding partners can agonize or antagonize target molecule biological activity. Also, test molecules that agonize or antagonize interactions between target molecules and binding partners can be useful as therapeutic molecules as they can up-regulate or down-regulated target molecule activity *in vivo* and thereby treat osteoarthritis.

[0156] Binding partners of target molecules can be identified by methods known in the art. For example, binding partners may be identified by lysing cells and analyzing cell lysates by electrophoretic techniques. Alternatively, a two-hybrid assay or three-hybrid assay can be utilized (*see, e.g.*, U.S. Patent No. 5,283,317; Zervos *et al.*, *Cell* 72:223-232 (1993); Madura *et al.*, *J. Biol. Chem.* 268: 12046-12054 (1993); Bartel *et al.*, *Biotechniques* 14: 920-924 (1993); Iwabuchi *et al.*, *Oncogene* 8: 1693-1696 (1993); and Brent WO94/10300). A two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. The assay often utilizes two different DNA constructs. In one construct, a *KIAA0296*, *chrom 4* or *SNW1* nucleic acid or other nucleic acid referenced in Table A (sometimes referred to as the “bait”) is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In another construct, a DNA sequence from a library of DNA sequences that encodes a potential binding partner (sometimes referred to as the “prey”) is fused to a gene that encodes an activation domain of the known transcription factor. Sometimes, a *KIAA0296*, *chrom 4* or *SNW1* nucleic acid or other nucleic acid referenced in Table A can be fused to the activation domain. If the “bait” and the “prey” molecules interact *in vivo*, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to identify the potential binding partner.

[0157] In an embodiment for identifying test molecules that antagonize or agonize complex formation between target molecules and binding partners, a reaction mixture containing the target molecule and the binding partner is prepared, under conditions and for a time sufficient to allow complex formation. The reaction mixture often is provided in the presence or absence of the test molecule. The test molecule can be included initially in the reaction mixture, or can be added at a time subsequent to the addition of the target molecule and its binding partner. Control reaction mixtures are incubated without the test molecule or with a placebo. Formation of any complexes between the target molecule and the

binding partner then is detected. Decreased formation of a complex in the reaction mixture containing test molecule as compared to in a control reaction mixture indicates that the molecule antagonizes target molecule/binding partner complex formation. Alternatively, increased formation of a complex in the reaction mixture containing test molecule as compared to in a control reaction mixture indicates that the molecule agonizes target molecule/binding partner complex formation. In another embodiment, complex formation of target molecule/binding partner can be compared to complex formation of mutant target molecule/binding partner (*e.g.*, amino acid modifications in a target polypeptide). Such a comparison can be important in those cases where it is desirable to identify test molecules that modulate interactions of mutant but not non-mutated target gene products.

[0158] The assays can be conducted in a heterogeneous or homogeneous format. In heterogeneous assays, target molecule and/or the binding partner are immobilized to a solid phase, and complexes are detected on the solid phase at the end of the reaction. In homogeneous assays, the entire reaction is carried out in a liquid phase. In either approach, the order of addition of reactants can be varied to obtain different information about the molecules being tested. For example, test compounds that agonize target molecule/binding partner interactions can be identified by conducting the reaction in the presence of the test molecule in a competition format. Alternatively, test molecules that agonize preformed complexes, *e.g.*, molecules with higher binding constants that displace one of the components from the complex, can be tested by adding the test compound to the reaction mixture after complexes have been formed.

[0159] In a heterogeneous assay embodiment, the target molecule or the binding partner is anchored onto a solid surface (*e.g.*, a microtiter plate), while the non-anchored species is labeled, either directly or indirectly. The anchored molecule can be immobilized by non-covalent or covalent attachments. Alternatively, an immobilized antibody specific for the molecule to be anchored can be used to anchor the molecule to the solid surface. The partner of the immobilized species is exposed to the coated surface with or without the test molecule. After the reaction is complete, unreacted components are removed (*e.g.*, by washing) such that a significant portion of any complexes formed will remain immobilized on the solid surface. Where the non-immobilized species is pre-labeled, the detection of label immobilized on the surface is indicative of complex. Where the non-immobilized species is not pre-labeled, an indirect label can be used to detect complexes anchored to the surface; *e.g.*, by using a labeled antibody specific for the initially non-immobilized species. Depending upon the order of addition of reaction components, test compounds that inhibit complex formation or that disrupt preformed complexes can be detected.

[0160] In another embodiment, the reaction can be conducted in a liquid phase in the presence or absence of test molecule, where the reaction products are separated from unreacted components, and the complexes are detected (*e.g.*, using an immobilized antibody specific for one of the binding components to anchor any complexes formed in solution, and a labeled antibody specific for the other partner to

detect anchored complexes). Again, depending upon the order of addition of reactants to the liquid phase, test compounds that inhibit complex or that disrupt preformed complexes can be identified.

[0161] In an alternate embodiment, a homogeneous assay can be utilized. For example, a preformed complex of the target gene product and the interactive cellular or extracellular binding partner product is prepared. One or both of the target molecule or binding partner is labeled, and the signal generated by the label(s) is quenched upon complex formation (*e.g.*, U.S. Patent No. 4,109,496 that utilizes this approach for immunoassays). Addition of a test molecule that competes with and displaces one of the species from the preformed complex will result in the generation of a signal above background. In this way, test substances that disrupt target molecule/binding partner complexes can be identified.

[0162] Candidate therapeutics for treating osteoarthritis are identified from a group of test molecules that interact with a target molecule. Test molecules are normally ranked according to the degree with which they modulate (*e.g.*, agonize or antagonize) a function associated with the target molecule (*e.g.*, DNA replication and/or processing, RNA transcription and/or processing, polypeptide production and/or processing, and/or biological function/activity), and then top ranking modulators are selected. Also, pharmacogenomic information described herein can determine the rank of a modulator. The top 10% of ranked test molecules often are selected for further testing as candidate therapeutics, and sometimes the top 15%, 20%, or 25% of ranked test molecules are selected for further testing as candidate therapeutics. Candidate therapeutics typically are formulated for administration to a subject.

Therapeutic Formulations

[0163] Formulations and pharmaceutical compositions typically include in combination with a pharmaceutically acceptable carrier one or more target molecule modulators. The modulator often is a test molecule identified as having an interaction with a target molecule by a screening method described above. The modulator may be a compound, an antisense nucleic acid, a ribozyme, an antibody, or a binding partner. Also, formulations may comprise a target polypeptide or fragment thereof in combination with a pharmaceutically acceptable carrier.

[0164] As used herein, the term “pharmaceutically acceptable carrier” includes solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Supplementary active compounds can also be incorporated into the compositions. Pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

[0165] A pharmaceutical composition typically is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application

can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerin, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

[0166] Oral compositions generally include an inert diluent or an edible carrier. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules, *e.g.*, gelatin capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

[0167] Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It should be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as mannitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

[0168] Sterile injectable solutions can be prepared by incorporating the active compound in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as

required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

[0169] For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

[0170] Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art. Molecules can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

[0171] In one embodiment, active molecules are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. Materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

[0172] It is advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier.

[0173] Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD₅₀ (the dose lethal to 50% of the population) and the ED₅₀ (the dose therapeutically effective in 50% of the

population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD_{50}/ED_{50} . Molecules which exhibit high therapeutic indices are preferred. While molecules that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

[0174] The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such molecules lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any molecules used in the methods described herein, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

[0175] As defined herein, a therapeutically effective amount of protein or polypeptide (*i.e.*, an effective dosage) ranges from about 0.001 to 30 mg/kg body weight, sometimes about 0.01 to 25 mg/kg body weight, often about 0.1 to 20 mg/kg body weight, and more often about 1 to 10 mg/kg, 2 to 9 mg/kg, 3 to 8 mg/kg, 4 to 7 mg/kg, or 5 to 6 mg/kg body weight. The protein or polypeptide can be administered one time per week for between about 1 to 10 weeks, sometimes between 2 to 8 weeks, often between about 3 to 7 weeks, and more often for about 4, 5, or 6 weeks. The skilled artisan will appreciate that certain factors may influence the dosage and timing required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of a protein, polypeptide, or antibody can include a single treatment or, preferably, can include a series of treatments.

[0176] With regard to polypeptide formulations, featured herein is a method for treating osteoarthritis in a subject, which comprises contacting one or more cells in the subject with a first polypeptide, where the subject comprises a second polypeptide having one or more polymorphic variations associated with cancer, and where the first polypeptide comprises fewer polymorphic variations associated with cancer than the second polypeptide. The first and second polypeptides are encoded by a nucleic acid which comprises a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A; a nucleotide sequence which encodes a polypeptide consisting of an amino acid sequence encoded by a nucleotide sequence referenced in SEQ ID NO: 1-3 or referenced in Table A; a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded

by a nucleotide sequence of SEQ ID NO: 1-3 or referenced in Table A and a nucleotide sequence 90% or more identical to a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A. The subject often is a human.

[0177] For antibodies, a dosage of 0.1 mg/kg of body weight (generally 10 mg/kg to 20 mg/kg) is often utilized. If the antibody is to act in the brain, a dosage of 50 mg/kg to 100 mg/kg is often appropriate. Generally, partially human antibodies and fully human antibodies have a longer half-life within the human body than other antibodies. Accordingly, lower dosages and less frequent administration is often possible. Modifications such as lipidation can be used to stabilize antibodies and to enhance uptake and tissue penetration (*e.g.*, into the brain). A method for lipidation of antibodies is described by Cruikshank *et al.*, *J. Acquired Immune Deficiency Syndromes and Human Retrovirology* 14:193 (1997).

[0178] Antibody conjugates can be used for modifying a given biological response, the drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a polypeptide such as tumor necrosis factor, alpha-interferon, beta-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophage colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors. Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

[0179] For compounds, exemplary doses include milligram or microgram amounts of the compound per kilogram of subject or sample weight, for example, about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram. It is understood that appropriate doses of a small molecule depend upon the potency of the small molecule with respect to the expression or activity to be modulated. When one or more of these small molecules is to be administered to an animal (*e.g.*, a human) in order to modulate expression or activity of a polypeptide or nucleic acid described herein, a physician, veterinarian, or researcher may, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

[0180] With regard to nucleic acid formulations, gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (*see, e.g.*, U.S. Patent 5,328,470) or by stereotactic injection (*see e.g.*, Chen *et al.*, (1994) *Proc. Natl. Acad. Sci. USA* 91:3054-3057). Pharmaceutical preparations of gene therapy vectors can include a gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells (*e.g.*, retroviral vectors) the pharmaceutical preparation can include one or more cells which produce the gene delivery system. Examples of gene delivery vectors are described herein.

Therapeutic Methods

[0181] A therapeutic formulation described above can be administered to a subject in need of a therapeutic for inducing a desired biological response. Therapeutic formulations can be administered by any of the paths described herein. With regard to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from pharmacogenomic analyses described herein.

[0182] As used herein, the term “treatment” is defined as the application or administration of a therapeutic formulation to a subject, or application or administration of a therapeutic agent to an isolated tissue or cell line from a subject with the purpose to cure, heal, alleviate, relieve, alter, remedy, ameliorate, improve or affect osteoarthritis, symptoms of osteoarthritis or a predisposition towards osteoarthritis. A therapeutic formulation includes, but is not limited to, small molecules, peptides, antibodies, ribozymes and antisense oligonucleotides. Administration of a therapeutic formulation can occur prior to the manifestation of symptoms characteristic of osteoarthritis, such that osteoarthritis is prevented or delayed in its progression. The appropriate therapeutic composition can be determined based on screening assays described herein.

[0183] As discussed, successful treatment of osteoarthritis can be brought about by techniques that serve to agonize target molecule expression or function, or alternatively, antagonize target molecule expression or function. These techniques include administration of modulators that include, but are not limited to, small organic or inorganic molecules; antibodies (including, for example, polyclonal, monoclonal, humanized, anti-idiotypic, chimeric or single chain antibodies, and Fab, F(ab')₂ and Fab expression library fragments, scFV molecules, and epitope-binding fragments thereof); and peptides, phosphopeptides, or polypeptides.

[0184] Further, antisense and ribozyme molecules that inhibit expression of the target gene can also be used to reduce the level of target gene expression, thus effectively reducing the level of target gene activity. Still further, triple helix molecules can be utilized in reducing the level of target gene activity. Antisense, ribozyme and triple helix molecules are discussed above. It is possible that the use of

antisense, ribozyme, and/or triple helix molecules to reduce or inhibit mutant gene expression can also reduce or inhibit the transcription (triple helix) and/or translation (antisense, ribozyme) of mRNA produced by normal target gene alleles, such that the concentration of normal target gene product present can be lower than is necessary for a normal phenotype. In such cases, nucleic acid molecules that encode and express target gene polypeptides exhibiting normal target gene activity can be introduced into cells via gene therapy method. Alternatively, in instances in that the target gene encodes an extracellular polypeptide, it can be preferable to co-administer normal target gene polypeptide into the cell or tissue in order to maintain the requisite level of cellular or tissue target gene activity.

[0185] Another method by which nucleic acid molecules may be utilized in treating or preventing osteoarthritis is use of aptamer molecules specific for target molecules. Aptamers are nucleic acid molecules having a tertiary structure which permits them to specifically bind to ligands (*see, e.g., Osborne, et al., Curr. Opin. Chem. Biol.* 1(1): 5-9 (1997); and Patel, D. J., *Curr. Opin. Chem. Biol. Jun;1(1): 32-46 (1997)*).

[0186] Yet another method of utilizing nucleic acid molecules for osteoarthritis treatment is gene therapy, which can also be referred to as allele therapy. Provided herein is a gene therapy method for treating osteoarthritis in a subject, which comprises contacting one or more cells in the subject or from the subject with a nucleic acid having a first nucleotide sequence (e.g., the first nucleotide sequence is identical to or substantially identical to a nucleotide sequence of SEQ ID NO: 1-3 or other nucleotide sequence referenced in Table A). Genomic DNA in the subject comprises a second nucleotide sequence having one or more polymorphic variations associated with osteoarthritis (e.g., the second nucleotide sequence is identical to or substantially identical to a nucleotide sequence of SEQ ID NO: 1-3 or other nucleotide sequence referenced in Table A). The first and second nucleotide sequences typically are substantially identical to one another, and the first nucleotide sequence comprises fewer polymorphic variations associated with osteoarthritis than the second nucleotide sequence. The first nucleotide sequence may comprise a gene sequence that encodes a full-length polypeptide or a fragment thereof. The subject is often a human. Allele therapy methods often are utilized in conjunction with a method of first determining whether a subject has genomic DNA that includes polymorphic variants associated with osteoarthritis.

[0187] In another allele therapy embodiment, provided herein is a method which comprises contacting one or more cells in the subject or from the subject with a polypeptide encoded by a nucleic acid having a first nucleotide sequence (e.g., the first nucleotide sequence is identical to or substantially identical to the nucleotide sequence of SEQ ID NO: 1-3 or other nucleotide sequence referenced in Table A). Genomic DNA in the subject comprises a second nucleotide sequence having one or more polymorphic variations associated with osteoarthritis (e.g., the second nucleotide sequence is identical to or substantially identical to a nucleotide sequence of SEQ ID NO: 1-3 or other nucleotide sequence

referenced in Table A). The first and second nucleotide sequences typically are substantially identical to one another, and the first nucleotide sequence comprises fewer polymorphic variations associated with osteoarthritis than the second nucleotide sequence. The first nucleotide sequence may comprise a gene sequence that encodes a full-length polypeptide or a fragment thereof. The subject is often a human.

[0188] For antibody-based therapies, antibodies can be generated that are both specific for target molecules and that reduce target molecule activity. Such antibodies may be administered in instances where antagonizing a target molecule function is appropriate for the treatment of osteoarthritis.

[0189] In circumstances where stimulating antibody production in an animal or a human subject by injection with a target molecule is harmful to the subject, it is possible to generate an immune response against the target molecule by use of anti-idiotypic antibodies (*see, e.g., Herlyn, Ann. Med.; 31(1): 66-78 (1999); and Bhattacharya-Chatterjee & Foon, Cancer Treat. Res.; 94: 51-68 (1998)*). Introducing an anti-idiotypic antibody to a mammal or human subject often stimulates production of anti-anti-idiotypic antibodies, which typically are specific to the target molecule. Vaccines directed to osteoarthritis also may be generated in this fashion.

[0190] In instances where the target molecule is intracellular and whole antibodies are used, internalizing antibodies may be preferred. Lipofectin or liposomes can be used to deliver the antibody or a fragment of the Fab region that binds to the target antigen into cells. Where fragments of the antibody are used, the smallest inhibitory fragment that binds to the target antigen is preferred. For example, peptides having an amino acid sequence corresponding to the Fv region of the antibody can be used. Alternatively, single chain neutralizing antibodies that bind to intracellular target antigens can also be administered. Such single chain antibodies can be administered, for example, by expressing nucleotide sequences encoding single-chain antibodies within the target cell population (*see, e.g., Marasco et al., Proc. Natl. Acad. Sci. USA 90: 7889-7893 (1993)*).

[0191] Modulators can be administered to a patient at therapeutically effective doses to treat osteoarthritis. A therapeutically effective dose refers to an amount of the modulator sufficient to result in amelioration of symptoms of osteoarthritis. Toxicity and therapeutic efficacy of modulators can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.,* for determining the LD₅₀ (the dose lethal to 50% of the population) and the ED₅₀ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD₅₀/ED₅₀. Modulators that exhibit large therapeutic indices are preferred. While modulators that exhibit toxic side effects can be used, care should be taken to design a delivery system that targets such molecules to the site of affected tissue in order to minimize potential damage to uninfected cells, thereby reducing side effects.

[0192] Data obtained from cell culture assays and animal studies can be used in formulating a range of dosages for use in humans. The dosage of such compounds lies preferably within a range of

circulating concentrations that include the ED_{50} with little or no toxicity. The dosage can vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the methods described herein, the therapeutically effective dose can be estimated initially from cell culture assays. A dose can be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} (*i.e.*, the concentration of the test compound that achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma can be measured, for example, by high performance liquid chromatography.

[0193] Another example of effective dose determination for an individual is the ability to directly assay levels of “free” and “bound” compound in the serum of the test subject. Such assays may utilize antibody mimics and/or “biosensors” that have been created through molecular imprinting techniques. Molecules that modulate target molecule activity are used as a template, or “imprinting molecule”, to spatially organize polymerizable monomers prior to their polymerization with catalytic reagents. The subsequent removal of the imprinted molecule leaves a polymer matrix which contains a repeated “negative image” of the compound and is able to selectively rebind the molecule under biological assay conditions. A detailed review of this technique can be seen in Ansell *et al.*, *Current Opinion in Biotechnology* 7: 89-94 (1996) and in Shea, *Trends in Polymer Science* 2: 166-173 (1994). Such “imprinted” affinity matrixes are amenable to ligand-binding assays, whereby the immobilized monoclonal antibody component is replaced by an appropriately imprinted matrix. An example of the use of such matrixes in this way can be seen in Vlatakis, *et al.*, *Nature* 361: 645-647 (1993). Through the use of isotope-labeling, the “free” concentration of compound which modulates target molecule expression or activity readily can be monitored and used in calculations of IC_{50} . Such “imprinted” affinity matrixes can also be designed to include fluorescent groups whose photon-emitting properties measurably change upon local and selective binding of target compound. These changes readily can be assayed in real time using appropriate fiberoptic devices, in turn allowing the dose in a test subject to be quickly optimized based on its individual IC_{50} . An example of such a “biosensor” is discussed in Kriz *et al.*, *Analytical Chemistry* 67: 2142-2144 (1995).

[0194] The examples set forth below are intended to illustrate but not limit the invention.

Examples

[0195] In the following studies a group of subjects was selected according to specific parameters pertaining to osteoarthritis. Nucleic acid samples obtained from individuals in the study group were subjected to genetic analyses that identified associations between osteoarthritis and certain polymorphic variants in human genomic DNA. The polymorphisms were genotyped again in two replication cohorts consisting of individuals selected for OA. In addition, SNPs proximal to the incident polymorphism in

the *KIAA0296* region, the *SNWI* region and in chromosome 4 were identified and allelotyped in OA case and control pools. Methods are described for producing target polypeptides encoded by the nucleic acids of Table A *in vitro* or *in vivo*, which can be utilized in methods that screen test molecules for those that interact with target polypeptides. Test molecules identified as being interactors with target polypeptides can be screened further as osteoarthritis therapeutics.

Example 1

Samples and Pooling Strategies

Sample Selection

[0196] Blood samples were collected from individuals diagnosed with knee osteoarthritis, which were referred to as case samples. Also, blood samples were collected from individuals not diagnosed with knee osteoarthritis as gender and age-matched controls. A database was created that listed all phenotypic trait information gathered from individuals for each case and control sample. Genomic DNA was extracted from each of the blood samples for genetic analyses.

DNA Extraction from Blood Samples

[0197] Six to ten milliliters of whole blood was transferred to a 50 ml tube containing 27 ml of red cell lysis solution (RCL). The tube was inverted until the contents were mixed. Each tube was incubated for 10 minutes at room temperature and inverted once during the incubation. The tubes were then centrifuged for 20 minutes at 3000 x g and the supernatant was carefully poured off. 100-200 µl of residual liquid was left in the tube and was pipetted repeatedly to resuspend the pellet in the residual supernatant. White cell lysis solution (WCL) was added to the tube and pipetted repeatedly until completely mixed. While no incubation was normally required, the solution was incubated at 37°C or room temperature if cell clumps were visible after mixing until the solution was homogeneous. 2 ml of protein precipitation was added to the cell lysate. The mixtures were vortexed vigorously at high speed for 20 sec to mix the protein precipitation solution uniformly with the cell lysate, and then centrifuged for 10 minutes at 3000 x g. The supernatant containing the DNA was then poured into a clean 15 ml tube, which contained 7 ml of 100% isopropanol. The samples were mixed by inverting the tubes gently until white threads of DNA were visible. Samples were centrifuged for 3 minutes at 2000 x g and the DNA was visible as a small white pellet. The supernatant was decanted and 5 ml of 70% ethanol was added to each tube. Each tube was inverted several times to wash the DNA pellet, and then centrifuged for 1 minute at 2000 x g. The ethanol was decanted and each tube was drained on clean absorbent paper. The DNA was dried in the tube by inversion for 10 minutes, and then 1000 µl of 1X TE was added. The size of each sample was estimated, and less TE buffer was added during the following DNA hydration step if

the sample was smaller. The DNA was allowed to rehydrate overnight at room temperature, and DNA samples were stored at 2-8°C.

[0198] DNA was quantified by placing samples on a hematology mixer for at least 1 hour. DNA was serially diluted (typically 1:80, 1:160, 1:320, and 1:640 dilutions) so that it would be within the measurable range of standards. 125 µl of diluted DNA was transferred to a clear U-bottom microtitre plate, and 125 µl of 1X TE buffer was transferred into each well using a multichannel pipette. The DNA and 1X TE were mixed by repeated pipetting at least 15 times, and then the plates were sealed. 50 µl of diluted DNA was added to wells A5-H12 of a black flat bottom microtitre plate. Standards were inverted six times to mix them, and then 50 µl of 1X TE buffer was pipetted into well A1, 1000 ng/ml of standard was pipetted into well A2, 500 ng/ml of standard was pipetted into well A3, and 250 ng/ml of standard was pipetted into well A4. PicoGreen (Molecular Probes, Eugene, Oregon) was thawed and freshly diluted 1:200 according to the number of plates that were being measured. PicoGreen was vortexed and then 50µl was pipetted into all wells of the black plate with the diluted DNA. DNA and PicoGreen were mixed by pipetting repeatedly at least 10 times with the multichannel pipette. The plate was placed into a Fluoroskan Ascent Machine (microplate fluorometer produced by Labsystems) and the samples were allowed to incubate for 3 minutes before the machine was run using filter pairs 485 nm excitation and 538 nm emission wavelengths. Samples having measured DNA concentrations of greater than 450 ng/µl were re-measured for confirmation. Samples having measured DNA concentrations of 20 ng/µl or less were re-measured for confirmation.

Pooling Strategies – Discovery Cohort

[0199] Samples were derived from the Nottingham knee OA family study (UK) where index cases were identified through a knee replacement registry. Siblings were approached and assessed with knee x-rays and assigned status as affected or unaffected. In all 1,157 individuals were available. In order to create same-sex pools of appropriate sizes, 335 unrelated female individuals with OA from the Nottingham OA sample were selected for the case pool. The control pool was made up of unrelated female individuals from the St. Thomas twin study (England) with normal knee x-rays and without other indications of OA, regardless of anatomical location, as well as lacking family history of OA. The St. Thomas twin study consists of Caucasian, female participants from the St. Thomas' Hospital, London, adult-twin registry, which is a voluntary registry of >4,000 twin pairs ranging from 18 to 76 years of age. The female case samples and female control samples are described further in Table 1 below.

[0200] A select set of samples from each group were utilized to generate pools, and one pool was created for each group. Each individual sample in a pool was represented by an equal amount of genomic DNA. For example, where 25 ng of genomic DNA was utilized in each PCR reaction and there

were 200 individuals in each pool, each individual would provide 125 pg of genomic DNA. Inclusion or exclusion of samples for a pool was based upon the following criteria: the sample was derived from an individual characterized as Caucasian; the sample was derived from an individual of British paternal and maternal descent; case samples were derived from individuals diagnosed with specific knee osteoarthritis (OA) and were recruited from an OA knee replacement clinic. Control samples were derived from individuals free of OA, family history of OA, and rheumatoid arthritis. Also, sufficient genomic DNA was extracted from each blood sample for all allelotyping and genotyping reactions performed during the study. Phenotype information from each individual was collected and included age of the individual, gender, family history of OA, general medical information (e.g., height, weight, thyroid disease, diabetes, psoriasis, hysterectomy), joint history (previous and current symptoms, joint-related operations, age at onset of symptoms, date of primary diagnosis, age of individual as of primary diagnosis and order of involvement), and knee-related findings (crepitus, restricted passive movement, bony swelling/deformity). Additional knee information included knee history, current symptoms, any major knee injury, meniscectomy, knee replacement surgery, age of surgery, and treatment history (including hormone replace therapy (HRT)). Samples that met these criteria were added to appropriate pools based on disease status.

[0201] The selection process yielded the pools set forth in Table 1, which were used in the studies that follow:

TABLE 1

	Female case	Female control
Pool size (Number)	335	335
Pool Criteria (ex: case/control)	control	case
Mean Age (ex: years)	57.21	69.95

Example 2

Association of Polymorphic Variants with Osteoarthritis

[0202] A whole-genome screen was performed to identify particular SNPs associated with occurrence of osteoarthritis. As described in Example 1, two sets of samples were utilized, which included samples from female individuals having knee osteoarthritis (osteoarthritis cases), and samples from female individuals not having knee osteoarthritis (female controls). The initial screen of each pool was performed in an allelotyping study, in which certain samples in each group were pooled. By pooling DNA from each group, an allele frequency for each SNP in each group was calculated. These allele

frequencies were then compared to one another. Particular SNPs were considered as being associated with osteoarthritis when allele frequency differences calculated between case and control pools were statistically significant. SNP disease association results obtained from the allelotyping study were then validated by genotyping each associated SNP across all samples from each pool. The results of the genotyping then were analyzed, allele frequencies for each group were calculated from the individual genotyping results, and a p-value was calculated to determine whether the case and control groups had statistically significant differences in allele frequencies for a particular SNP. When the genotyping results agreed with the original allelotyping results, the SNP disease association was considered validated at the genetic level.

SNP Panel Used for Genetic Analyses

[0203] A whole-genome SNP screen began with an initial screen of approximately 25,000 SNPs over each set of disease and control samples using a pooling approach. The pools studied in the screen are described in Example 1. The SNPs analyzed in this study were part of a set of 25,488 SNPs confirmed as being statistically polymorphic as each is characterized as having a minor allele frequency of greater than 10%. The SNPs in the set reside in genes or in close proximity to genes, and many reside in gene exons. Specifically, SNPs in the set are located in exons, introns, and within 5,000 base-pairs upstream of a transcription start site of a gene. In addition, SNPs were selected according to the following criteria: they are located in ESTs; they are located in Locuslink or Ensembl genes; and they are located in Genomatix promoter predictions. SNPs in the set were also selected on the basis of even spacing across the genome, as depicted in Table 2.

[0204] A case-control study design using a whole genome association strategy involving approximately 28,000 single nucleotide polymorphisms (SNPs) was employed. Approximately 25,000 SNPs were evenly spaced in gene-based regions of the human genome with a median inter-marker distance of about 40,000 base pairs. Additionally, approximately 3,000 SNPs causing amino acid substitutions in genes described in the literature as candidates for various diseases were used. The case-control study samples were of female Caucasian origin (British paternal and maternal descent) 670 individuals were equally distributed in two groups: female controls and female cases. The whole genome association approach was first conducted on 2 DNA pools representing the 2 groups. Significant markers were confirmed by individual genotyping.

TABLE 2

<u>General Statistics</u>		<u>Spacing Statistics</u>	
Total # of SNPs	25,488	Median	37,058 bp
# of Exonic SNPs	>4,335 (17%)	Minimum*	1,000 bp
# SNPs with refSNP ID	20,776 (81%)	Maximum*	3,000,000 bp
Gene Coverage	>10,000	Mean	122,412 bp
Chromosome Coverage	All	Std Deviation	373,325 bp
		<i>*Excludes outliers</i>	

Allelotyping and Genotyping Results

[0205] The genetic studies summarized above and described in more detail below identified allelic variants associated with osteoarthritis, which are summarized in Table A.

Assay for Verifying, Allelotyping, and Genotyping SNPs

[0206] A MassARRAY™ system (Sequenom, Inc.) was utilized to perform SNP genotyping in a high-throughput fashion. This genotyping platform was complemented by a homogeneous, single-tube assay method (hME™ or homogeneous MassEXTEND™ (Sequenom, Inc.)) in which two genotyping primers anneal to and amplify a genomic target surrounding a polymorphic site of interest. A third primer (the MassEXTEND™ primer), which is complementary to the amplified target up to but not including the polymorphism, was then enzymatically extended one or a few bases through the polymorphic site and then terminated.

[0207] For each polymorphism, SpectroDESIGNER™ software (Sequenom, Inc.) was used to generate a set of PCR primers and a MassEXTEND™ primer which were used to genotype the polymorphism. Other primer design software could be used or one of ordinary skill in the art could manually design primers based on his or her knowledge of the relevant factors and considerations in designing such primers. Table 3 shows PCR primers and Table 4 shows extension primers used for analyzing polymorphisms. The initial PCR amplification reaction was performed in a 5 µl total volume containing 1X PCR buffer with 1.5 mM MgCl₂ (Qiagen), 200 µM each of dATP, dGTP, dCTP, dTTP (Gibco-BRL), 2.5 ng of genomic DNA, 0.1 units of HotStar DNA polymerase (Qiagen), and 200 nM each of forward and reverse PCR primers specific for the polymorphic region of interest.

TABLE 3: PCR Primers

SNP Reference	Forward PCR primer	Reverse PCR primer
rs552	ACGTTGGATGGACTGAGGTAGATGATGC	ACGTTGGATGGCTTTCTTCCCTTGGTTTC
rs12904	ACGTTGGATGGAACCACTCCCACCACAG	ACGTTGGATGGGTGGGGATGGCACTGTC
rs2282146	ACGTTGGATGTCCCACGAGGACCTGGAG	ACGTTGGATGTCGTTTGGGTGGCCGGG

SNP Reference	Forward PCR primer	Reverse PCR primer
rs734784	ACGTTGGATGTCGGGATGTCTCCAGAGATG	ACGTTGGATGGCAACCACCAAGAGTTTGAG
rs1042164	ACGTTGGATGTTTCTTCCAGACGGGCTTTC	ACGTTGGATGCAAAGTCAGCCGCAAACGAC
rs749670	ACGTTGGATGTCTCATCTGTGTGCCCATTG	ACGTTGGATGATGAGGGTGAAAGGCAGGAG
rs955592	ACGTTGGATGTTCCCATTCCTTCTGGGCTC	ACGTTGGATGTCTCAGAGGGTCTCCTTTTC
rs1143016	ACGTTGGATGTTGTCCAGCAGGTAGGGCAG	ACGTTGGATGACCCATCGCGGATACATGTG
rs755248	ACGTTGGATGGGTCTCTGCTGAGGAAGTGG	ACGTTGGATGACACTCACTACGGGGCCAG
rs1055055	ACGTTGGATGTTGTGCTTGCTGAGGAATCC	ACGTTGGATGGTTGCAGAGAGCGTCTATAC
rs835409	ACGTTGGATGTCCTGTTGGCTTTTGCAGAC	ACGTTGGATGACTGCTCATGGTGGTTGAAC
rs927663	ACGTTGGATGTTGACTGGTTGCCCAAAC	ACGTTGGATGAAGAATCTTCAGTGCCAGCC
rs8162	ACGTTGGATGCTTCATCCAGAACCTCCAGG	ACGTTGGATGTGCATATGGCTTGTGAGAGC
rs831038	ACGTTGGATGTGAAAGAGCTGCCTTCTTTC	ACGTTGGATGAAATGACACTCACGGTAAGC
rs33079	ACGTTGGATGTTATTTTCATTGGCCAAGCCC	ACGTTGGATGGTGTTCACCTGTTTCATGCAC
rs1710880	ACGTTGGATGCGAAGGCAGAGAATAAACTG	ACGTTGGATGAACTCTGTGGTTTAAAGAAAG
rs1078153	ACGTTGGATGTCCTGCGTGTAACTGAGAGG	ACGTTGGATGAACATACACACAGTGCGAGC
rs799570	ACGTTGGATGATGCATATGGGCAGGTTGCC	ACGTTGGATGCCAGGAAAGCATCCTCAGAC
rs1282730	ACGTTGGATGTCCTTTGACTTACTGTGCTG	ACGTTGGATGAGAAAAGAGGTTGTGTACAG
rs1518875	ACGTTGGATGAGAATGCGTTCAATGCCTGC	ACGTTGGATGAGCGAAAAGCTCTGCCATTG
rs1568694	ACGTTGGATGGTTCATTCACTTATGGACGG	ACGTTGGATGTGATAGGAGGGAGCCATCTC
rs905042	ACGTTGGATGTAACAATGGTAAGGGCCAAG	ACGTTGGATGGGTCCATAATGGTCATTGTG
rs1957723	ACGTTGGATGTACTCACTTGTGTACTGCTC	ACGTTGGATGGCTGCAGCGTCACATTAATC
rs794018	ACGTTGGATGGGATGATGATGAAATGACTG	ACGTTGGATGGCTCTAGTTAGATGAGTCTC
rs707723	ACGTTGGATGTGTGGCTGAAGTTTGCTCTG	ACGTTGGATGCACACACAAACCTTGAAGAG
rs893861	ACGTTGGATGGAGGCATGTACACAAAACCTG	ACGTTGGATGGCTCAGCACTGTAATAGTTG
rs1914903	ACGTTGGATGTGCGTCAAGTTGAAGTCCTC	ACGTTGGATGAGGGTAGTGAGTTCACATGC
rs2062232	ACGTTGGATGTCTGCTCAGATAAAGTCTG	ACGTTGGATGGCGGTAGTTTTCCCTAAACC
rs26609	ACGTTGGATGCAAGGGAGATCAGAAACATC	ACGTTGGATGAATTCATTGTTCTTGATGGC
rs1370987	ACGTTGGATGATACTTTGGATGTCTGGTGG	ACGTTGGATGGGTCTTTGGTCACAACATATC
rs1012414	ACGTTGGATGACTTGGAAAGTCAGTCTGGG	ACGTTGGATGGAAACCGAGAAATGGCTATG
rs435903	ACGTTGGATGGGCATAAGTTAGAGACAACC	ACGTTGGATGGGCTATGTTATGCTGCTGTG
rs1248	ACGTTGGATGGAGATTGTGCATTTTGGCAAG	ACGTTGGATGCAGACACCATCTTAACCAAG
rs703508	ACGTTGGATGAGCTCTGTGGCCTCTTTTGG	ACGTTGGATGTACTCACAGTCTTCCCGCGG
rs226465	ACGTTGGATGAATTTTGACCCCTGCCAACC	ACGTTGGATGTATGTGAAAGAGGCGGTGAAG
rs241448	ACGTTGGATGCAAGCTGCAGCAAGCTTGCC	ACGTTGGATGTGAGAAAGAGGCGCCGATATC
rs763155	ACGTTGGATGGGGAAACCCAAAATAGTGTG	ACGTTGGATGTCACAGGAGAGTAATGCCTC
rs1040461	ACGTTGGATGACATCTGGTGGAAGTCACTC	ACGTTGGATGGGTCTTTGTTTGGTGGGTC
rs462832	ACGTTGGATGCACTTTCTCTGTAATATTTG	ACGTTGGATGTGAGACAACAAAAATTTGCC
rs804194	ACGTTGGATGTAATCCGGTGGCAGATCAAG	ACGTTGGATGGAAATTCATGTGCTGACGGG
rs1022646	ACGTTGGATGACTGTACCTAATCATCCTG	ACGTTGGATGGACTATGTTGGAGTTCAGAG
rs1569112	ACGTTGGATGTATGATCTGCCTGTGGAGA	ACGTTGGATGACCATCCTCACACCCATCCA
rs805623	ACGTTGGATGATCAACCACTCATACACTGC	ACGTTGGATGCACAGAAACAGCTGGATTGC
rs1019850	ACGTTGGATGTTTTACTCCAGGAAGCCACC	ACGTTGGATGAGCAGGGAGAAATTGTTCCAG
rs1599931	ACGTTGGATGTCAAACCTTCTGTAGACG	ACGTTGGATGTGAACATAGTAGGCGCTCAG
AA	ACGTTGGATGTAGGAGTGCTCGTATTTTGG	ACGTTGGATGCTGGGAACAGCTTTTGATCC
rs279941	ACGTTGGATGCATAGGGAACACCGAGAATG	ACGTTGGATGGGTGTGTCATCTATGGGCTAG
rs1062230	ACGTTGGATGAAACTCCTTTCCCTCTCAAG	ACGTTGGATGGGCCCATCAGTCTATAGTTT
rs1859911	ACGTTGGATGCTGTTTTCCGAGCATCTAC	ACGTTGGATGCCTCTTGATATGAGATAGG
rs1477261	ACGTTGGATGCAGGGTTATGTGGTATTATC	ACGTTGGATGGGGAAAGTAAAAGATAAGAG
rs1191119	ACGTTGGATGACTCTCAGGGTGATTATCTG	ACGTTGGATGTGTAAGATCTGGCACTGTC
rs657780	ACGTTGGATGTTTAAAGAGCCGCCAAGGAG	ACGTTGGATGCCCATTTTCAGACCACTTGG
rs1393890	ACGTTGGATGGTCTGATTATCTTCTGCCG	ACGTTGGATGGGTACCTTTATCCTTGCTTC
rs1478714	ACGTTGGATGAATAATTTGCTGACACCCCC	ACGTTGGATGGGAGTCCAGAGGTTAAACAG

SNP Reference	Forward PCR primer	Reverse PCR primer
rs868213	ACGTTGGATGTGTCAGAACTGGGCACATTC	ACGTTGGATGAGGGATAGGGATCAGGAATG
rs690115	ACGTTGGATGAATAGCCAAGGCCGTGTGGG	ACGTTGGATGCACCTGGGAGATAGCAGGG
rs1465501	ACGTTGGATGTCAGGAATTGTTACCTGGAC	ACGTTGGATGCCCTCATCTAGACACTTTTG
rs899173	ACGTTGGATGAGTGCCACATCACTCTTG	ACGTTGGATGTTCTGCTCCACTACAGTCTC
rs10477	ACGTTGGATGGGGGCTACGTGGAAGTTACC	ACGTTGGATGATGGCAATCAAGAGAGTCTAA
rs926393	ACGTTGGATGAGATCAGCCCAGGAAATGTG	ACGTTGGATGTGTTGGAGAAGGTTCCACC
rs465271	ACGTTGGATGAATCACAGCTCATGGCTCAC	ACGTTGGATGATGGTAGTGTGCACCTATGG
rs13847	ACGTTGGATGCGCCCGTAGTGATAAGCAC	ACGTTGGATGCAGGACAGGGCAGAGTGAG
rs738658	ACGTTGGATGGATGGTATGTGTGCATCAGG	ACGTTGGATGCTTCCAAGAGATGGCGTTC

[0208] Samples were incubated at 95°C for 15 minutes, followed by 45 cycles of 95°C for 20 seconds, 56°C for 30 seconds, and 72°C for 1 minute, finishing with a 3 minute final extension at 72°C. Following amplification, shrimp alkaline phosphatase (SAP) (0.3 units in a 2 µl volume) (Amersham Pharmacia) was added to each reaction (total reaction volume was 7 µl) to remove any residual dNTPs that were not consumed in the PCR step. Samples were incubated for 20 minutes at 37°C, followed by 5 minutes at 85°C to denature the SAP.

[0209] Once the SAP reaction was complete, a primer extension reaction was initiated by adding a polymorphism-specific MassEXTEND™ primer cocktail to each sample. Each MassEXTEND™ cocktail included a specific combination of dideoxynucleotides (ddNTPs) and deoxynucleotides (dNTPs) used to distinguish polymorphic alleles from one another. Methods for verifying, allelotyping and genotyping SNPs are disclosed, for example, in U.S. Pat. No. 6,258,538, the content of which is hereby incorporated by reference. In Table 4, ddNTPs are shown and the fourth nucleotide not shown is the dNTP.

TABLE 4: Extension Primers

SNP Reference	Extend Probe	Termination Mix
rs552	TGATGCTGTTGTCAGATACC	ACT
rs12904	AGCCTCAAAACGGGTCA	ACT
rs2282146	GGACCTGGAGCCCCACC	ACG
rs734784	GCCTCCGACACCCCATCAA	ACG
rs1042164	CTTGCTCGGGACCACTCCA	ACT
rs749670	GGTGGTGGGCATCCCTTTC	ACG
rs955592	TTGGGCTCTGACCACCTCT	ACT
rs1143016	ATGCAGCGTCACCAGCAC	ACT
rs755248	TGAGGAAGTGGCAGGTGTG	ACG
rs1055055	CCCAGTTCAGGCTCACTTTC	ACT
rs835409	TTGCAGACCAGCCAATTAAGAA	ACT
rs927663	GGTTGCCCCAACTCCCTT	ACT
rs8162	AACACAGAGCAAAGCACC	ACT
rs831038	CGTTATAGTAAAGGAAAGGCAG	ACG

SNP Reference	Extend Probe	Termination Mix
rs33079	CCCATCACCTGGAGCTTTG	ACG
rs1710880	CTGTATTATGTTTCCCCTTGG	CGT
rs1078153	GCCGGCACCGTCAGAAAC	CGT
rs799570	GCAGTTCCTAGAAGACAGCT	ACT
rs1282730	TGCTGGCCCAACTTTTGTCT	ACG
rs1518875	CTGCAATGTTTCCAAACCCC	ACT
rs1568694	AGTTATGGACGGAAGAAGGG	ACG
rs905042	GGTAAGGGCCAAGTGAGTG	CGT
rs1957723	AGCATGGCATAGGCACTGG	ACG
rs794018	AAATGACTGAAAATGTGACTATA	ACG
rs707723	CCTGAGGTATATTCAATA	ACG
rs893861	CATGTACACAAAACGTTAAGTAA	ACG
rs1914903	TCCCCATAGATGGACCTGC	ACG
rs2062232	GCTGAAGACAAGGATTAGGTT	ACG
rs26609	GAGATCAGAAACATCACCTTG	CGT
rs1370987	TTTGGGAGTTACTGCCTTAGAA	ACT
rs1012414	ACTAGGAACCAGAATATGAGCATC	ACG
rs435903	AAGCTAACAATGGAATAATGGC	ACG
rs1248	GTGCATTTTGGCAAGAATATATG	CGT
rs703508	GGGGTCCAGGCAGAAAGAAAC	ACG
rs226465	CCTCTTCCCCTCCTCCCT	ACT
rs241448	GCAGAAGCTTGCCCAGCTC	ACG
rs763155	GCAGCCTGCAAGTGAGTGA	ACT
rs1040461	AAGTCACTCCGGTCAGAATTCA	ACT
rs462832	ATAAGAACTCTTTAGATCCCAACA	CGT
rs804194	GATCAAGGCTGATCTCGCC	ACT
rs1022646	CCTAATCATCCTGCCACCC	ACT
rs1569112	ACCAGGCCGCATGGGCTG	ACG
rs805623	CTGTGTTCAAATAAGGCAACC	ACT
rs1019850	AGGAAGCCACCAGCTAATAC	CGT
rs1599931	CTGAGGCCGGGAGGGATT	ACT
AA	TAGTTTTAAATTCTGCACA	ACT
rs279941	AACACCGAGAATGAAAACATC	ACT
rs1062230	ATGCTGGTTCTGTCCAA	ACG
rs1859911	TCCGAGCATCTACATGCTCA	ACT
rs1477261	AGGAGGAGCCCAAATATGAAA	CGT
rs1191119	GTCTTTTGTTAAGTGGGGAACCC	ACG
rs657780	CGCCAAGGAGTTTCCCACA	ACT
rs1393890	CTGCCGTACCTGGCAAGC	ACT
rs1478714	CCCCGAGGGGACAGTCCA	ACG
rs868213	GGGCACATTCTTGAGGAGGT	ACG
rs690115	AGCCGAGGGAGCTGACCCTG	ACG
rs1465501	TCCAGGAGCCCTCAGAATG	ACT
rs899173	CCTCTGGCAAAGTGTGGAGC	ACG
rs10477	AGTACGATATCAAAGATC	ACG
rs926393	CAGGAAATGTGCTTTTCGAGTTCC	ACG
rs465271	GGCTCAAGGGATCCTCCCA	ACG
rs13847	AAGCACACCGGCACGAAC	ACT
rs738658	GAGGCATTTTCATTAATGCATG	CGT

[0210] The MassEXTEND™ reaction was performed in a total volume of 9 µl, with the addition of 1X ThermoSequenase buffer, 0.576 units of ThermoSequenase (Amersham Pharmacia), 600 nM MassEXTEND™ primer, 2 mM of ddATP and/or ddCTP and/or ddGTP and/or ddTTP, and 2 mM of dATP or dCTP or dGTP or dTTP. The deoxy nucleotide (dNTP) used in the assay normally was complementary to the nucleotide at the polymorphic site in the amplicon. Samples were incubated at 94°C for 2 minutes, followed by 55 cycles of 5 seconds at 94°C, 5 seconds at 52°C, and 5 seconds at 72°C.

[0211] Following incubation, samples were desalted by adding 16 µl of water (total reaction volume was 25 µl), 3 mg of SpectroCLEAN™ sample cleaning beads (Sequenom, Inc.) and allowed to incubate for 3 minutes with rotation. Samples were then robotically dispensed using a piezoelectric dispensing device (SpectroJET™ (Sequenom, Inc.)) onto either 96-spot or 384-spot silicon chips containing a matrix that crystallized each sample (SpectroCHIP™ (Sequenom, Inc.)). Subsequently, MALDI-TOF mass spectrometry (Biflex and Autoflex MALDI-TOF mass spectrometers (Bruker Daltonics) can be used) and SpectroTYPER RT™ software (Sequenom, Inc.) were used to analyze and interpret the SNP genotype for each sample.

Genetic Analysis

[0212] Minor allelic frequencies for the polymorphisms set forth in Table A were verified as being 10% or greater using the extension assay described above in a group of samples isolated from 92 individuals originating from the state of Utah in the United States, Venezuela and France (Coriell cell repositories).

[0213] Genotyping results are shown for female pools in Table 5. In Table 5, “AF” refers to allelic frequency; and “F case” and “F control” refer to female case and female control groups, respectively.

TABLE 5: Genotyping Results

SNP Reference	AF F case	AF F control	p-value
rs552	A = 0.190 G = 0.810	A = 0.123 G = 0.877	0.0011
rs12904	A = 0.455 G = 0.545	A = 0.375 G = 0.625	0.0012
rs2282146	C = 0.906 T = 0.094	C = 0.939 T = 0.061	0.0105
rs734784	G = 0.483 A = 0.517	G = 0.416 A = 0.584	0.0052
rs1042164	T = 0.233 C = 0.767	T = 0.159 C = 0.841	0.0002
rs749670	C = 0.342 T = 0.658	C = 0.419 T = 0.581	0.0038
rs955592	T = 0.045 C = 0.955	T = 0.076 C = 0.924	0.0177

SNP Reference	AF F case	AF F control	p-value
rs1143016	T = 0.093 C = 0.907	T = 0.054 C = 0.946	0.0071
rs755248	G = 0.146 A = 0.854	G = 0.069 A = 0.931	0.0000
rs1055055	A = 0.432 G = 0.568	A = 0.355 G = 0.645	0.0046
rs835409	T = 0.620 G = 0.380	T = 0.681 G = 0.319	0.0222
rs927663	T = 0.301 G = 0.699	T = 0.358 G = 0.642	0.0289
rs8162	A = 0.591 G = 0.409	A = 0.657 G = 0.343	0.0149
rs831038	C = 0.617 T = 0.383	C = 0.666 T = 0.334	0.0359
rs33079	G = 0.823 A = 0.177	G = 0.881 A = 0.119	0.0013
rs1710880	C = 0.303 A = 0.697	C = 0.371 A = 0.629	0.0129
rs1078153	T = 0.818 A = 0.182	T = 0.875 A = 0.125	0.0039
rs799570	A = 0.675 G = 0.325	A = 0.740 G = 0.260	0.0100
rs1282730	G = 0.086 A = 0.914	G = 0.127 A = 0.873	0.0150
rs1518875	T = 0.033 C = 0.967	T = 0.055 C = 0.945	0.0508
rs1568694	G = 0.045 A = 0.955	G = 0.081 A = 0.919	0.0064
rs905042	A = 0.832 T = 0.168	A = 0.769 T = 0.231	0.0047
rs1957723	G = 0.778 A = 0.222	G = 0.839 A = 0.161	0.0048
rs794018	G = 0.273 A = 0.727	G = 0.220 A = 0.780	0.0034
rs707723	C = 0.759 T = 0.241	C = 0.811 T = 0.189	0.0195
rs893861	G = 0.246 A = 0.754	G = 0.196 A = 0.804	0.0251
rs1914903	G = 0.861 A = 0.139	G = 0.910 A = 0.090	0.0055
rs2062232	C = 0.064 T = 0.936	C = 0.117 T = 0.883	0.0012
rs26609	A = 0.777 T = 0.223	A = 0.840 T = 0.160	0.0039
rs1370987	A = 0.422 G = 0.578	A = 0.341 G = 0.659	0.0007
rs1012414	G = 0.876 A = 0.124	G = 0.833 A = 0.167	0.0289
rs435903	G = 0.766 A = 0.234	G = 0.685 A = 0.315	0.0013
rs1248	T = 0.668 A = 0.332	T = 0.593 A = 0.407	0.0014
rs703508	G = 0.875 A = 0.125	G = 0.910 A = 0.090	0.0375
rs226465	G = 0.094 C = 0.906	G = 0.129 C = 0.871	0.0454
rs241448	C = 0.294 T = 0.706	C = 0.212 T = 0.788	0.0010
rs763155	A = 0.160	A = 0.114	0.0140

SNP Reference	AF F case	AF F control	p-value
	C = 0.840	C = 0.886	
rs1040461	T = 0.069 C = 0.931	T = 0.098 C = 0.902	0.0281
rs462832	A = 0.218 T = 0.782	A = 0.145 T = 0.855	0.0008
rs804194	T = 0.583 C = 0.417	T = 0.679 C = 0.321	0.0004
rs1022646	A = 0.169 G = 0.831	A = 0.103 G = 0.897	0.0007
rs1569112	G = 0.853 A = 0.147	G = 0.812 A = 0.188	0.0468
rs805623	A = 0.097 G = 0.903	A = 0.140 G = 0.860	0.0143
rs1019850	A = 0.330 T = 0.670	A = 0.240 T = 0.760	0.0005
rs1599931	A = 0.581 G = 0.419	A = 0.659 G = 0.341	0.0037
AA	A = 0.506 G = 0.494	A = 0.577 G = 0.423	0.0102
rs279941	T = 0.100 G = 0.900	T = 0.138 G = 0.862	0.0324
rs1062230	C = 0.778 T = 0.222	C = 0.717 T = 0.283	0.0109
rs1859911	T = 0.295 C = 0.705	T = 0.243 C = 0.757	0.0328
rs1477261	T = 0.861 A = 0.139	T = 0.809 A = 0.191	0.0105
rs1191119	G = 0.121 A = 0.879	G = 0.078 A = 0.922	0.0079
rs657780	A = 0.674 G = 0.326	A = 0.583 G = 0.417	0.0009
rs1393890	G = 0.639 C = 0.361	G = 0.724 C = 0.276	0.0014
rs1478714	G = 0.331 A = 0.669	G = 0.269 A = 0.731	0.0136
rs868213	C = 0.078 T = 0.922	C = 0.044 T = 0.956	0.0083
rs690115	G = 0.839 A = 0.161	G = 0.784 A = 0.216	0.0111
rs1465501	A = 0.846 G = 0.154	A = 0.903 G = 0.097	0.0020
rs899173	C = 0.895 T = 0.105	C = 0.858 T = 0.142	0.0408
rs10477	C = 0.087 T = 0.913	C = 0.146 T = 0.854	0.0010
rs926393	C = 0.715 T = 0.285	C = 0.647 T = 0.353	0.0082
rs465271	C = 0.194 T = 0.806	C = 0.130 T = 0.870	0.0019
rs13847	A = 0.111 G = 0.889	A = 0.163 G = 0.837	0.0056
rs738658	C = 0.898 A = 0.102	C = 0.855 A = 0.145	0.0183

[0214] All of the single marker alleles set forth in Table A were considered validated, since the genotyping data agreed with the allelotyping data and each SNP significantly associated with

osteoarthritis. Particularly significant associations with osteoarthritis are indicated by a calculated p-value of less than 0.05 for genotype results.

Example 3

Association of Polymorphic Variants with Osteoarthritis in Replication Cohorts

[0215] The single marker polymorphisms set forth in Table A were genotyped again in two replication cohorts consisting of individuals selected for OA.

Sample Selection and Pooling Strategies – Replication Sample 1

[0216] A second case control sample (replication sample #1) was created by using 100 Caucasian female cases from Chingford, UK, and 148 unrelated female cases from the St. Thomas twin study. Cases were defined as having Kellgren-Lawrence (KL) scores of at least 2 in at least one knee x-ray. In addition, 199 male knee replacement cases from Nottingham were included. (For a cohort description, see the Nottingham description provided in Example 1). The control pool was made up of unrelated female individuals from the St. Thomas twin study (England) with normal knee x-rays and without other indications of OA, regardless of anatomical location, as well as lacking family history of OA. The St. Thomas twin study consists of Caucasian, female participants from the St. Thomas' Hospital, London, adult-twin registry, which is a voluntary registry of >4,000 twin pairs ranging from 18 to 76 years of age. The replication sample 1 cohort was used to replicate the initial results. Table 6 below summarizes the selected phenotype data collected from the case and control individuals.

TABLE 6

Phenotype	Female cases (n=248): median (range)/ (n,%)	Male cases (n=199): median (range)/ (n,%)	Female controls (n=313): mean (range)/ (n,%)
Age	59 (39- 73)	66 (45- 73)	55 (50- 72)
Height (cm)	162 (141- 178)	175 (152- 198)	162 (141- 176)
Weight (kg)	68 (51- 123)	86 (62- 127)	64 (40- 111)
Body mass index (kg/m ²)	26 (18- 44)	29 (21- 41)	24 (18- 46)
Kellgren-Lawrence* left knee	0 (63, 26%), 1 (20, 8%), 2 (105, 43%), 3 (58, 23%), 4 (1, 0%)	NA	NA
Kellgren-Lawrence* right knee	0 (43, 7%), 1 (18, 7%), 2 (127, 52%), 3 (57, 23%), 4 (1, 0%)	NA	NA
KL* >2 both knees	No (145, 59%), Yes (101, 41%)	NA	NA

Phenotype	Female cases (n=248): median (range)/ (n,%)	Male cases (n=199): median (range)/ (n,%)	Female controls (n=313): mean (range)/ (n,%)
KL* >2 either knee	No (0, 0%), Yes (248, 100%)	NA	NA

* 0: normal, 1: doubtful, 2: definite osteophyte (bony protuberance), 3: joint space narrowing (with or without osteophyte), 4: joint deformity

Sample Selection and Pooling Strategies – Replication Sample 2

[0217] A third case control sample (replication sample #2) was created by using individuals with symptoms of OA from Newfoundland, Canada. These individuals were recruited and examined by rheumatologists. Affected joints were x-rayed and a final diagnosis of definite or probable OA was made according to American College of Rheumatology criteria by a single rheumatologist to avoid any inter-examiner diagnosis variability. Controls were recruited from volunteers without any symptoms from the musculoskeletal system based on a normal joint exam performed by a rheumatologist. Only cases with a diagnosis of definite OA were included in the study. Only individuals of Caucasian origin were included. The cases consisted of 228 individuals with definite knee OA, 106 individuals with definite hip OA, and 74 individuals with hip OA.

TABLE 7

Phenotype	Case	Control
Age at Visit	62.7	52.5
Sex (Female/Male)	227/119	174/101
Knee OA Xray: No	35% (120)	80% (16)
Unknown	1% (4)	0% (0)
Yes	64% (221)	20% (4)
Hip OA Xray: No	63% (215)	80% (16)
Unknown	2% (7)	0% (0)
Yes	35% (121)	20% (4)

Assay for Verifying, Allelotyping, and Genotyping SNPs

[0218] Genotyping of the replication cohorts described in Tables 6 and 7 was performed using the same methods used for the original genotyping, as described herein. A MassARRAY™ system (Sequenom, Inc.) was utilized to perform SNP genotyping in a high-throughput fashion. This genotyping platform was complemented by a homogeneous, single-tube assay method (hME™ or homogeneous MassEXTEND™ (Sequenom, Inc.)) in which two genotyping primers anneal to and amplify a genomic

target surrounding a polymorphic site of interest. A third primer (the MassEXTEND™ primer), which is complementary to the amplified target up to but not including the polymorphism, was then enzymatically extended one or a few bases through the polymorphic site and then terminated.

[0219] For each polymorphism, SpectroDESIGNER™ software (Sequenom, Inc.) was used to generate a set of PCR primers and a MassEXTEND™ primer which were used to genotype the polymorphism. Other primer design software could be used or one of ordinary skill in the art could manually design primers based on his or her knowledge of the relevant factors and considerations in designing such primers. Table 3 shows PCR primers and Table 4 shows extension probes used for analyzing (*e.g.*, genotyping) polymorphisms in the replication cohorts. The initial PCR amplification reaction was performed in a 5 µl total volume containing 1X PCR buffer with 1.5 mM MgCl₂ (Qiagen), 200 µM each of dATP, dGTP, dCTP, dTTP (Gibco-BRL), 2.5 ng of genomic DNA, 0.1 units of HotStar DNA polymerase (Qiagen), and 200 nM each of forward and reverse PCR primers specific for the polymorphic region of interest.

[0220] Samples were incubated at 95°C for 15 minutes, followed by 45 cycles of 95°C for 20 seconds, 56°C for 30 seconds, and 72°C for 1 minute, finishing with a 3 minute final extension at 72°C. Following amplification, shrimp alkaline phosphatase (SAP) (0.3 units in a 2 µl volume) (Amersham Pharmacia) was added to each reaction (total reaction volume was 7 µl) to remove any residual dNTPs that were not consumed in the PCR step. Samples were incubated for 20 minutes at 37°C, followed by 5 minutes at 85°C to denature the SAP.

[0221] Once the SAP reaction was complete, a primer extension reaction was initiated by adding a polymorphism-specific MassEXTEND™ primer cocktail to each sample. Each MassEXTEND™ cocktail included a specific combination of dideoxynucleotides (ddNTPs) and deoxynucleotides (dNTPs) used to distinguish polymorphic alleles from one another. Methods for verifying, allelotyping and genotyping SNPs are disclosed, for example, in U.S. Pat. No. 6,258,538, the content of which is hereby incorporated by reference. In Table 7, ddNTPs are shown and the fourth nucleotide not shown is the dNTP.

[0222] The MassEXTEND™ reaction was performed in a total volume of 9 µl, with the addition of 1X ThermoSequenase buffer, 0.576 units of ThermoSequenase (Amersham Pharmacia), 600 nM MassEXTEND™ primer, 2 mM of ddATP and/or ddCTP and/or ddGTP and/or ddTTP, and 2 mM of dATP or dCTP or dGTP or dTTP. The deoxy nucleotide (dNTP) used in the assay normally was complementary to the nucleotide at the polymorphic site in the amplicon. Samples were incubated at 94°C for 2 minutes, followed by 55 cycles of 5 seconds at 94°C, 5 seconds at 52°C, and 5 seconds at 72°C.

[0223] Following incubation, samples were desalted by adding 16 µl of water (total reaction volume was 25 µl), 3 mg of SpectroCLEAN™ sample cleaning beads (Sequenom, Inc.) and allowed to incubate for 3 minutes with rotation. Samples were then robotically dispensed using a piezoelectric dispensing device (SpectroJET™ (Sequenom, Inc.)) onto either 96-spot or 384-spot silicon chips containing a matrix that crystallized each sample (SpectroCHIP™ (Sequenom, Inc.)). Subsequently, MALDI-TOF mass spectrometry (Biflex and Autoflex MALDI-TOF mass spectrometers (Bruker Daltonics) can be used) and SpectroTYPER RT™ software (Sequenom, Inc.) were used to analyze and interpret the SNP genotype for each sample.

Genetic Analysis

[0224] Genotyping results for replication cohorts #1 and #2 are provided in Tables 8 and 9, respectively.

TABLE 8

rsID	Replication #1 (Mixed Male/Female cases and Female controls)				Meta-analysis Disc. + Rep #1
	AF OA Con	AF OA Cas	Delta	P-value	P-value
rs552	0.87	0.85	0.02	0.344	0.0300
rs12904	0.57	0.57	0.00	0.936	0.2700
rs2282146	0.08	0.1	-0.02	0.342	0.0190
rs734784	0.52	0.54	-0.02	0.451	0.7200
rs1042164	0.79	0.82	-0.03	0.161	0.9100
rs749670	0.62	0.66	-0.04	0.173	0.0019
rs955592	0.93	0.94	-0.01	0.521	0.0600
rs1143016	0.93	0.93	0.00	0.869	NA
rs755248	0.9	0.89	0.01	0.544	0.1600
rs1055055	0.64	0.64	0.00	0.947	0.3300
rs835409	0.34	0.35	-0.01	0.715	0.1300
rs927663	0.64	0.65	-0.01	0.611	0.0690
rs831038	0.35	0.37	-0.02	0.399	NA
rs33079	0.14	0.14	0.00	0.995	0.3100
rs1710880	0.66	0.62	0.04	0.087	0.9000
rs799570	0.29	0.29	0.00	0.903	0.2500
rs1282730	0.88	0.87	0.01	0.751	0.4800
rs1568694	0.93	0.94	0.00	0.928	0.2600
rs905042	0.21	0.2	0.01	0.829	0.2200
rs1957723	0.13	0.16	-0.03	0.124	0.0009
rs794018	0.74	0.72	0.02	0.518	0.0710
rs707723	0.18	0.19	-0.01	0.658	0.0650
rs1914903	0.15	0.14	0.01	0.605	0.5500
rs2062232	0.91	0.91	0.00	0.788	0.2100
rs26609	0.16	0.19	-0.02	0.226	0.0032

rsID	Replication #1 (Mixed Male/Female cases and Female controls)				Meta-analysis Disc. + Rep #1
	AF OA Con	AF OA Cas	Delta	P-value	P-value
rs1370987	0.63	0.63	-0.01	0.857	0.3900
rs1012414	0.12	0.13	-0.01	0.669	0.5600
rs435903	0.27	0.27	0.00	0.950	0.2800
rs1248	0.36	0.36	0.00	0.917	0.2400
rs703508	0.11	0.12	-0.01	0.558	0.0660
rs226465	0.87	0.88	-0.01	0.436	0.0500
rs241448	0.74	0.75	-0.01	0.805	0.4100
rs763155	0.86	0.88	-0.02	0.273	0.8800
rs1040461	0.92	0.92	0.00	0.826	NA
rs1022646	0.85	0.87	-0.02	0.219	0.8200
rs1569112	0.16	0.18	-0.02	0.402	0.8800
rs805623	0.87	0.88	-0.01	0.460	0.0370
rs1019850	0.69	0.7	0.00	0.890	0.3700
AA	0.47	0.48	-0.01	0.681	0.1200
rs279941	0.87	0.89	-0.01	0.400	0.0340
rs1062230	0.26	0.26	0.00	0.896	0.4200
rs1859911	0.71	0.75	-0.04	0.128	0.9000
rs1477261	0.16	0.16	0.00	0.986	0.3000
rs1191119	0.89	0.88	0.01	0.569	0.1200
rs1393890	0.29	0.31	-0.02	0.527	0.1400
rs1478714	0.69	0.67	0.03	0.300	0.0140
rs868213	0.92	0.93	-0.01	0.455	0.7000
rs690115	0.2	0.21	-0.01	0.729	0.4900
rs1465501	0.11	0.1	0.01	0.718	0.5600
rs899173	0.1	0.11	0.00	0.924	0.3300
rs10477	0.89	0.88	0.01	0.691	0.4700
rs926393	0.3	0.31	-0.01	0.830	0.4200
rs465271	0.86	0.85	0.01	0.516	0.0660
rs13847	0.86	0.85	0.01	0.547	0.5900
rs738658	0.14	0.15	-0.01	0.536	0.6700

TABLE 9

rsID	Replication #2 (Newfoundland) (Male/Female cases and controls)				Meta-analysis Disc. + Rep #2
	AF OA Con	AF OA Cas	Delta	P-value	Not Done
rs552	0.85	0.86	-0.014	0.496	
rs12904	0.58	0.57	0.011	0.719	
rs2282146	0.08	0.08	0.002	0.876	
rs734784	0.53	0.54	-0.003	0.907	
rs1042164	0.83	0.80	0.026	0.248	
rs749670	0.66	0.62	0.036	0.208	
rs955592	0.95	0.92	0.033	0.027	

rsID	Replication #2 (Newfoundland) (Male/Female cases and controls)				Meta-analysis
	AF OA Con	AF OA Cas	Delta	P-value	Disc. + Rep #2 Not Done
rs1143016	0.96	0.94	0.015	0.236	
rs755248	0.89	0.90	-0.009	0.608	
rs1055055	0.64	0.61	0.034	0.249	
rs835409	0.36	0.31	0.047	0.101	
rs927663	0.67	0.68	-0.013	0.631	
rs831038	0.34	0.35	-0.014	0.612	
rs33079	0.17	0.19	-0.019	0.417	
rs1710880	0.64	0.62	0.029	0.309	
rs799570	0.35	0.30	0.058	0.033	
rs1282730	0.89	0.89	-0.001	0.982	
rs1568694	0.95	0.94	0.009	0.518	
rs905042	0.19	0.20	-0.002	0.933	
rs1957723	0.18	0.20	-0.017	0.454	
rs794018	0.73	0.72	0.015	0.586	
rs707723	0.20	0.21	-0.007	0.759	
rs1914903	0.14	0.16	-0.022	0.285	
rs2062232	0.92	0.91	0.008	0.632	
rs26609	0.19	0.18	0.005	0.827	
rs1370987	0.59	0.61	-0.023	0.423	
rs1012414	0.15	0.14	0.008	0.679	
rs435903	0.24	0.26	-0.026	0.316	
rs1248	0.33	0.38	-0.051	0.078	
rs703508	0.10	0.11	-0.002	0.916	
rs226465	0.89	0.89	-0.007	0.699	
rs241448	0.76	0.77	-0.007	0.778	
rs763155	0.89	0.84	0.049	0.016	
rs1040461	0.91	0.91	0.001	0.948	
rs1022646	0.86	0.86	-0.001	0.974	
rs1569112	0.16	0.17	-0.016	0.446	
rs805623	0.89	0.87	0.022	0.256	
rs1019850	0.71	0.69	0.026	0.341	
AA	0.48	0.44	0.035	0.234	
rs279941	0.91	0.87	0.037	0.047	
rs1062230	0.23	0.22	0.011	0.653	
rs1859911	0.72	0.71	0.015	0.560	
rs1477261	0.17	0.14	0.031	0.143	
rs1191119	0.86	0.88	-0.017	0.377	
rs1393890	0.30	0.28	0.017	0.516	
rs1478714	0.68	0.70	-0.025	0.358	
rs868213	0.91	0.93	-0.019	0.260	
rs690115	0.19	0.18	0.005	0.811	
rs1465501	0.10	0.12	-0.020	0.282	
rs899173	0.14	0.12	0.020	0.319	
rs10477	0.86	0.88	-0.016	0.442	
rs926393	0.37	0.32	0.042	0.137	

rsID	Replication #2 (Newfoundland) (Male/Female cases and controls)				Meta-analysis Disc. + Rep #2 Not Done
	AF OA Con	AF OA Cas	Delta	P-value	
rs465271	0.87	0.85	0.023	0.263	
rs13847	0.84	0.85	-0.012	0.582	
rs738658	0.18	0.15	0.021	0.340	

[0225] To combine the evidence for association from multiple sample collections, a meta-analysis procedure was employed. The allele frequencies were compared between cases and controls within the discovery sample, as well as within the replication cohort #1 using the DerSimonian-Laird approach (DerSimonian, R. and N. Laird. 1986. Meta-analysis in clinical trials. Control Clin Trials 7: 177-188.)

[0226] The absence of a statistically significant association in one or more of the replication cohorts should not be interpreted as minimizing the value of the original finding. There are many reasons why a biologically derived association identified in a sample from one population would not replicate in a sample from another population. The most important reason is differences in population history. Due to bottlenecks and founder effects, there may be common disease predisposing alleles present in one population that are relatively rare in another, leading to a lack of association in the candidate region. Also, because common diseases such as arthritis-related disorders are the result of susceptibilities in many genes and many environmental risk factors, differences in population-specific genetic and environmental backgrounds could mask the effects of a biologically relevant allele. For these and other reasons, statistically strong results in the original, discovery sample that did not replicate in one or more of the replication samples may be further evaluated in additional replication cohorts and experimental systems.

Example 4

KIAA0296 Region Proximal SNPs

[0227] SNP rs749670 is associated with osteoarthritis and is described in Table A. It lies within the KIAA0296 gene and codes for a G327E amino acid change. The thymine allele of SNP rs749670 is associated with osteoarthritis (see Table 5) and codes for glutamic acid. KIAA0296 shares homology with C2H2-type Zn-finger protein and is likely a novel transcription factor. One-hundred one additional allelic variants proximal to rs749670 were identified and subsequently allelotyped in osteoarthritis case and control sample sets as described in Examples 1 and 2. The polymorphic variants are set forth in Table 10. The chromosome positions provided in column four of Table 10 are based on Genome “Build 34” of NCBI’s GenBank.

TABLE 10

dbSNP rs#	Chromosome	Position in SEQ ID NO: 1	Chromosome Position	Allele Variants
rs7500176	16	247	31077197	a/g
rs6565212	16	1535	31078485	c/t
rs8054046	16	2386	31079336	c/t
rs8056842	16	6440	31083390	c/t
rs732173	16	9133	31086083	g/t
rs732172	16	9143	31086093	a/g
rs7188557	16	9471	31086421	a/t
rs2288004	16	13150	31090100	c/g
rs4337310	16	13717	31090667	c/t
rs2016554	16	14466	31091416	a/g
rs6565213	16	15769	31092719	a/c
rs7204762	16	16870	31093820	a/g
rs4889529	16	18545	31095495	c/t
rs6565214	16	18749	31095699	c/t
rs7499674	16	19123	31096073	g/t
rs6565215	16	20736	31097686	a/g
rs1023623	16	21038	31097988	c/t
rs1023624	16	21046	31097996	c/t
rs1023625	16	21050	31098000	c/t
rs1549297	16	21056	31098006	a/t
rs3084894	16	21706	31098656	-/acc
rs8048228	16	23170	31100120	a/g
rs7405432	16	25028	31101978	a/t
rs8054249	16	27871	31104821	a/g
rs8061047	16	28070	31105020	c/t
rs7187220	16	31717	31108667	a/g
rs8046978	16	32019	31108969	a/g
rs2288003	16	32318	31109268	a/g
rs7196421	16	33080	31110030	a/g
rs7196431	16	33101	31110051	a/g
rs7203158	16	34236	31111186	a/g
rs2303223	16	34285	31111235	c/t
rs2032917	16	34818	31111768	c/g
rs8044134	16	35168	31112118	c/g
rs4889531	16	37981	31114931	c/t
rs4889532	16	38113	31115063	c/g
rs4889533	16	38117	31115067	c/t
rs881929	16	38481	31115431	g/t
rs8047104	16	38615	31115565	c/g
rs8047803	16	38944	31115894	a/c
rs4644874	16	39288	31116238	a/c
rs2359673	16	41385	31118335	c/t
rs4435271	16	42136	31119086	a/t
rs7197717	16	42185	31119135	a/c
rs2359674	16	42353	31119303	a/g
rs6565217	16	42434	31119384	a/g

dbSNP rs#	Chromosome	Position in SEQ ID NO: 1	Chromosome Position	Allele Variants
rs2303222	16	44580	31121530	a/g
rs4889615	16	44675	31121625	a/t
rs4624197	16	45739	31122689	g/t
rs3751853	16	46439	31123389	c/t
rs749671	16	47457	31124407	c/t
rs749670	16	47735	31124685	c/t
rs3751855	16	50319	31127269	c/t
rs3751856	16	50708	31127658	a/g
rs7196726	16	51185	31128135	a/g
rs889550	16	53002	31129952	a/g
rs750952	16	53064	31130014	c/t
rs2077633	16	53637	31130587	a/g
rs7199949	16	55274	31132224	c/g
rs2032916	16	55825	31132775	c/t
rs4468641	16	55986	31132936	a/c
rs4889535	16	56684	31133634	c/g
rs4316775	16	57653	31134603	c/t
rs4313819	16	57659	31134609	c/g
rs6565218	16	57692	31134642	g/t
rs4318224	16	57775	31134725	c/t
rs1046030	16	61313	31138263	c/t
rs7294	16	61431	31138381	a/g
rs7200749	16	61699	31138649	a/g
rs2359612	16	62906	31139856	a/g
rs8050894	16	63619	31140569	c/g
rs2884737	16	64664	31141614	a/c
rs1895514	16	68452	31145402	g/t
rs8060209	16	69665	31146615	c/t
rs8060217	16	69681	31146631	c/t
rs7196161	16	70091	31147041	a/g
rs8062336	16	74637	31151587	a/g
rs8043778	16	74760	31151710	a/g
rs2032915	16	76523	31153473	a/g
rs4889616	16	78559	31155509	c/g
rs1045564	16	79549	31156499	a/c
rs2303221	16	79882	31156832	c/t
rs1549296	16	81339	31158289	a/g
rs889555	16	81681	31158631	c/t
rs5816521	16	81696	31158646	-/g
rs749767	16	83517	31160467	c/t
rs2884738	16	85431	31162381	a/c
rs2052581	16	86332	31163282	c/t
rs4889617	16	87358	31164308	a/g
rs4889619	16	87725	31164675	c/t
rs1978487	16	89052	31166002	a/g
rs1978486	16	90020	31166970	a/g
rs1978485	16	90231	31167181	a/g

dbSNP rs#	Chromosome	Position in SEQ ID NO: 1	Chromosome Position	Allele Variants
rs4889620	16	90284	31167234	a/g
rs4889621	16	90447	31167397	c/t
rs3214477	16	90601	31167551	-/g
rs4527034	16	90724	31167674	a/g
rs1060506	16	92559	31169509	c/t
rs7200125	16	95176	31172126	a/g
rs6565219	16	95195	31172145	c/t
rs889548	16	96822	31173772	a/g

Assay for Verifying and Allelotyping SNPs

[0228] The methods used to verify and allelotype the 101 proximal SNPs of Table 10 are the same methods described in Examples 1 and 2 herein. The primers and probes used in these assays are provided in Table 11 and Table 12, respectively.

TABLE 11

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs7500176	ACGTTGGATGACAGTGGCTCATGCCTGTAA	ACGTTGGATGTTTCACCATATTGGCCAGGC
rs6565212	ACGTTGGATGTTAGGAAGGATGTGGAAGGG	ACGTTGGATGGACCTGACCTCAAAGAGAAG
rs8054046	ACGTTGGATGCACTGAAGTTTAGAGCAGCC	ACGTTGGATGTGCACAGTGGGTAACTGTAG
rs8056842	ACGTTGGATGATGAGGTTTCACCTTGTTGG	ACGTTGGATGATCATAGCACTTTGCGAGGC
rs732173	ACGTTGGATGAGACCAGGCTCAGTCCAAAC	ACGTTGGATGTGGCCAAACCTGGAAGACAC
rs732172	ACGTTGGATGCTCAGTCCAAACTGCCAGAC	ACGTTGGATGCATGGCCAAACCTGGAAGAC
rs7188557	ACGTTGGATGAACATCTGTACAAGGCTGGG	ACGTTGGATGATTGGCTGTAGCATGACTGA
rs2288004	ACGTTGGATGAAAGACACTGGAAGGCTGTG	ACGTTGGATGAGAGAAGGTGGAGCTCTTTC
rs4337310	ACGTTGGATGAGGGAAGAGATGTACACAGG	ACGTTGGATGTTTGGAGCAGATCTGGTAGG
rs2016554	ACGTTGGATGAAGCAATCCTCCACCTCAG	ACGTTGGATGCAAGAGCAAACTCCCTCTC
rs6565213	ACGTTGGATGAGATGGAGTCTCACTCCATC	ACGTTGGATGTGAGGCAGGAGAATCGCTTG
rs7204762	ACGTTGGATGAGTGGCTCACACCTGTAATC	ACGTTGGATGGCTGGTCTTGAACCTCTGAC
rs4889529	ACGTTGGATGCAAGCAATCCTTGCCCTAAG	ACGTTGGATGGGTGGTTCACATCTGCAATC
rs6565214	ACGTTGGATGTGATCTCGGCTCACTGCAAG	ACGTTGGATGAAAATTAGCCGGGCATGGTG
rs7499674	ACGTTGGATGAACTAGGGAACCTTCCAC	ACGTTGGATGTGGGCCCCACTAAGTCTAAA
rs6565215	ACGTTGGATGAGACGGAAAGTTCCAGCTTG	ACGTTGGATGTGGGACCACTCTGTTCTATG
rs1023623	ACGTTGGATGACAGAGCAAGACTCCATCTC	ACGTTGGATGTCCTCTTCAGAGCTGTTAC
rs1023624	ACGTTGGATGTGACAGAGCAAGACTCCATC	ACGTTGGATGGTCTTAACCACTGAGCCTAT
rs1023625	ACGTTGGATGTGGTGACAGAGCAAGACTCC	ACGTTGGATGTCAGTCTTAACCACTGAGC
rs1549297	ACGTTGGATGTTGCATTGATCCGAGATCGC	ACGTTGGATGTCAGTCTTAACCACTGAGC
rs3084894	ACGTTGGATGTCCCAGGTTCAAGCGATTCT	ACGTTGGATGCCATGAAACCCCATCTCTAC
rs8048228	ACGTTGGATGAATTGCTTGAACCTGGGAGG	ACGTTGGATGTTGACAGTCTCCCTCTATC
rs7405432	ACGTTGGATGAGATCATGCCACTGCACTAC	ACGTTGGATGCACTGCACTTGGCCTAATTG
rs8054249	ACGTTGGATGATCTCCTGACCTCATGATCT	ACGTTGGATGTAATCAAACACCAGGCTGGG
rs8061047	ACGTTGGATGATGATCACAGCTCACTGCAG	ACGTTGGATGCTCCCTGCCTCTACAAAAAG
rs7187220	ACGTTGGATGAAGGAGACCTTCTCCACAAT	ACGTTGGATGCCGGTCAGAGAAGCTCTTGC

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs8046978	ACGTTGGATGTGCACAGGAGCTGGTGGTG	ACGTTGGATGATCACACCACCTGACTCCGG
rs2288003	ACGTTGGATGACCGGCCGTTCAAGTGCCTG	ACGTTGGATGAGAGTGCACCAGCGCGTGC
rs7196421	ACGTTGGATGTTACGCCATTCTCCTGCCT	ACGTTGGATGAAATTAGCCAGGCGTGGTGG
rs7196431	ACGTTGGATGAGATCTCGGCTCACTGCAAG	ACGTTGGATGATGTAGTCCCAGCTACTCGG
rs7203158	ACGTTGGATGAAGCCTATGCGGAGCTCAAG	ACGTTGGATGATTGGCTGCAGCAACGCTGT
rs2303223	ACGTTGGATGACCCTCACCGCTCATGGTTG	ACGTTGGATGTGCGGCCCTACAGCTGTGA
rs2032917	ACGTTGGATGCCTGGGCGCGTTTGAAATG	ACGTTGGATGAGCCCCCGGCTACAAGCGCT
rs8044134	ACGTTGGATGACTAAGAAAGGAGGCTGAGG	ACGTTGGATGACAGTGTTTGGAAAAGCCCG
rs4889531	ACGTTGGATGATTCTCACCCTCACTCTGTC	ACGTTGGATGGACCGTGTGTAATGTACTGC
rs4889532	ACGTTGGATGGGGACAAGAATCCCTATCTC	ACGTTGGATGTAGAGCCAGACACATTGCTG
rs4889533	ACGTTGGATGCTCTGTAAAGTAGGGACAAG	ACGTTGGATGTAGAGCCAGACACATTGCTG
rs881929	ACGTTGGATGTTGACCCAGTGTTCTGAGC	ACGTTGGATGCCAGCTACCTGGTGTCTAAC
rs8047104	ACGTTGGATGGTGGGATGTTAGACAGAGAC	ACGTTGGATGTGCCAGGTTGGTCTCAGCAT
rs8047803	ACGTTGGATGAAAGTGCTGGGATTACAGGC	ACGTTGGATGAAATACAGATTCTGAGGCC
rs4644874	ACGTTGGATGAGTCTTGCTATGTTGCCTGG	ACGTTGGATGTAATCCAGCACTTTGGGAG
rs2359673	ACGTTGGATGGTGTGATGTCAGTTCACTGC	ACGTTGGATGATCCCAAATACTTGGGAGGC
rs4435271	ACGTTGGATGACAGTGGTCTCAAGAACTCC	ACGTTGGATGTGGCTCATGCCTGTAATCAC
rs7197717	ACGTTGGATGTGTGATTACAGGCATGAGCC	ACGTTGGATGGCTTGCAAGGAGTATTGTCC
rs2359674	ACGTTGGATGGCCTAGCAGTTCATTATGAG	ACGTTGGATGCCCTTGCTCCAAATACAGTC
rs6565217	ACGTTGGATGAAGAACGCTAATCCTACTGG	ACGTTGGATGTGGAGACAAGGCCCTTTATGG
rs2303222	ACGTTGGATGTTGGGAAAAGTCTCCAGAG	ACGTTGGATGGCGCAGAAAGGGAGAAAAAG
rs4889615	ACGTTGGATGTAAGTTCTAGGTCTGCACGG	ACGTTGGATGATGCACCGGAACGATTCTAG
rs4624197	ACGTTGGATGCGCTAAGAGAGTCTTTGGG	ACGTTGGATGCAGAGCGAGACTCCATCTCA
rs3751853	ACGTTGGATGTCCCCTAGGCTTAAGTCATC	ACGTTGGATGGGTCTGTGATCAGAAGTAGG
rs749671	ACGTTGGATGTGACTACATTTGTACCGCCG	ACGTTGGATGTCAAGTAGTGAACCTCACAGG
rs749670	ACGTTGGATGTCTCATCTGTGTGCCATTG	ACGTTGGATGATGAGGGTGAAAGGCAGGAG
rs3751855	ACGTTGGATGAAGAAGAGGTGTGGGAGGAG	ACGTTGGATGTCAGAGCTGGCTTCAGTCTG
rs3751856	ACGTTGGATGAGCTGTACTGGCCCGTCTCG	ACGTTGGATGCAGTGCGGGCGGACCTATC
rs7196726	ACGTTGGATGGACCTAGTTAGGAAGTGAAG	ACGTTGGATGTCAGGGCAGCAAGCTCAGAAG
rs889550	ACGTTGGATGTCCACCCAGCACTGCTGGA	ACGTTGGATGCAGGTCTCTGCTGAGGGAAC
rs750952	ACGTTGGATGTTCCCTCAGCAGGACCTGG	ACGTTGGATGGGTGGCCACTAGATGGAATG
rs2077633	ACGTTGGATGTTTCTCAGGAGTAGTTCGGG	ACGTTGGATGAAAGAAGCCAGATCTGGGTC
rs7199949	ACGTTGGATGTCCCCTCAGGCAGGTGGT	ACGTTGGATGCAGCCTGTGACACTGGGAG
rs2032916	ACGTTGGATGGTTCCCCTCATTACTGAAGG	ACGTTGGATGTGCCACTTGCTGTAGTTAC
rs4468641	ACGTTGGATGATGAGTCAGGAATACGGGAG	ACGTTGGATGAATGCCCTACTTGTCACCTC
rs4889535	ACGTTGGATGCTATGGCAGACACCTCTGA	ACGTTGGATGGAAGAGAAGGAGCAGAAAGG
rs4316775	ACGTTGGATGAGTAGCTCACGCTTGTAATC	ACGTTGGATGCTATGTTGCACAGGCTAGTC
rs4313819	ACGTTGGATGTGCACAGGCTAGTCTTGAAC	ACGTTGGATGAGTAGCTCACGCTTGTAATC
rs6565218	ACGTTGGATGTTAAAGTCACAGACTGAGGC	ACGTTGGATGTTGAACTCTTGGGCTCAAGC
rs4318224	ACGTTGGATGTCACTCTGTGACTTTAAGCG	ACGTTGGATGACCACCTTTCATGGTAGAAG
rs1046030	ACGTTGGATGGTCTCCAAAGCTCTTCCATT	ACGTTGGATGGATTGATCTAAGAACTTTA
rs7294	ACGTTGGATGGCACTGGGTGTAAAAAGAG	ACGTTGGATGTTCTAGATTACCCCTCCTC
rs7200749	ACGTTGGATGGAGCACGAAGAAGGATCC	ACGTTGGATGTCTGTCTGATGCTGCTGAG
rs2359612	ACGTTGGATGAAATCGGCCAAGTCTGAACC	ACGTTGGATGTCCAGAGAAGGCATCACTGA
rs8050894	ACGTTGGATGAATCTTGGTGATCCACACAG	ACGTTGGATGTAGTTACCTCCACATCCC
rs2884737	ACGTTGGATGTCAATTATGCTAACGCCTGGC	ACGTTGGATGTTGACGATGGTCTCAAGGAC
rs1895514	ACGTTGGATGCAATCTCAGCTCACTGCAAC	ACGTTGGATGTAATCCAGCTACTTGGGAG

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs8060209	ACGTTGGATGGGTCAGGAGTTTAAGACAAG	ACGTTGGATGCCATGCCCGGCTAATTTTTG
rs8060217	ACGTTGGATGTGAGTAGCTGGGATTACAGG	ACGTTGGATGAGACAAGCTTGGCCAACATG
rs7196161	ACGTTGGATGGTGTGTTTTAGTAGAGACGGG	ACGTTGGATGATCCCAGCACTTTAGGAAGC
rs8062336	ACGTTGGATGTGCTCCCCACATCTCAGACG	ACGTTGGATGAAGCGAGGAGCGCCTCTTC
rs8043778	ACGTTGGATGTTCTCCTCACTTCTCAGACGGG	ACGTTGGATGATCGTCTGAGATGTGGGGAG
rs2032915	ACGTTGGATGATTCCCACCCGTTCTTTCCC	ACGTTGGATGTTCCCGCTCCCTTTTACCAC
rs4889616	ACGTTGGATGGAACCAAGAACTGGAAGGAG	ACGTTGGATGTGTAAAGCGCACAGATCACG
rs1045564	ACGTTGGATGTGTCAGCATCCTCGACGCAC	ACGTTGGATGACCCAGGCGACCCAAAATGG
rs2303221	ACGTTGGATGAGAACCCCAACACTCTCCC	ACGTTGGATGAGCGGAGAAGGTGCGCAAG
rs1549296	ACGTTGGATGATGCTGCTGAACCTCCTAAC	ACGTTGGATGAGCAGGGTTTCTCAACCATG
rs889555	ACGTTGGATGAGACCAGTAGGTACAAGCAC	ACGTTGGATGTCAAGAATGCCATGAGGTGG
rs5816521	ACGTTGGATGATTGTGGCTCTATGCAGAGG	ACGTTGGATGTCAAGAATGCCATGAGGTGG
rs749767	ACGTTGGATGCTGATAGAAAGGACCAAGGA	ACGTTGGATGCTGGAGTTCTGATTCAAGTC
rs2884738	ACGTTGGATGAGAACTGCTTGAACCCAGGA	ACGTTGGATGATGGAGTCTTGTGTGTCGG
rs2052581	ACGTTGGATGTGGGACATGCGGATATGGAG	ACGTTGGATGGAGGGTTCTGTGAGAGTCAG
rs4889617	ACGTTGGATGCAGAGCGAGACTCCATCTCA	ACGTTGGATGACACTCGCGCTGGCCTAATG
rs4889619	ACGTTGGATGAAAATTAGATGGGCGTGGTG	ACGTTGGATGATCTCGGCTCACTGCAACCT
rs1978487	ACGTTGGATGTCCCTTCTCTATGTTCTGCTG	ACGTTGGATGATGGAGGAAGACAGAGAGAG
rs1978486	ACGTTGGATGTACCTAGGGTCACAGATTG	ACGTTGGATGGGGTATGTGGTAAAATGAGC
rs1978485	ACGTTGGATGTCAAGCAATTTTCTGCCTC	ACGTTGGATGCCATCTGTACCAAAAAGACG
rs4889620	ACGTTGGATGTGGCAAACCCCATCTGTAC	ACGTTGGATGAGTAGTTGGGATTACAGGTG
rs4889621	ACGTTGGATGTACTCAATCACTGCCACAAC	ACGTTGGATGGCCAGTTATTTTCTCATTG
rs3214477	ACGTTGGATGACTCGAGACTGGATCACTTC	ACGTTGGATGCCTTTTGTTCAGCCTTACC
rs4527034	ACGTTGGATGAAGTATGGGCCATAAGAGTG	ACGTTGGATGTATGTACACTACGTGGGCTG
rs1060506	ACGTTGGATGATCAGGAGTGCAAACCAGAG	ACGTTGGATGGGATGAAGCTGCAATAGCTG
rs7200125	ACGTTGGATGATTTGCCATTGCACTCCAG	ACGTTGGATGTACAGGCATGAGCCATAGCC
rs6565219	ACGTTGGATGCTTGGCCTCTCAAAGTGCTG	ACGTTGGATGAGGGCGAGGCTCCATTTCAA
rs889548	ACGTTGGATGCTGGCCAAGTCCTAATACAG	ACGTTGGATGCCCAATTCAGAGATGTCAG

TABLE 12

dbSNP rs#	Extend Primer	Term Mix
rs7500176	GATCACGAGGTCAGGAGTTC	ACT
rs6565212	GCTGGAAAAGTGTGAGGGT	ACT
rs8054046	TTTAGAGCAGCCGATACCCA	ACG
rs8056842	GCTGGTCTCGAACTCCTGA	ACG
rs732173	GCTCAGTCCAACTGCCAG	CGT
rs732172	ACTGCCAGACTCCCGCCA	ACG
rs7188557	CCTGGCCCTGGTTGTGAGT	CGT
rs2288004	CGGCAGATCCAGTGTGTC	ACT
rs4337310	CACGGAATCTCCAGTGCAC	ACT
rs2016554	GGCACGTACCACTGACATG	ACG
rs6565213	GCAGTGGCGCAATCTTGAC	ACT
rs7204762	CCCAGCACTTTGGGAGGC	ACG

dbSNP rs#	Extend Primer	Term Mix
rs4889529	CTCAAGTGATCCTCCTGCCT	ACG
rs6565214	GAGTAGCTGGGACTACAGG	ACG
rs7499674	GTTCTTCTCAACATCTGCCCA	ACT
rs6565215	TTTCCTTCAGACAGGGCTCT	ACT
rs1023623	GACTCCATCTCAAAAAAAAAAAAAA	ACT
rs1023624	GAGCAAGACTCCATCTCAAAAA	ACT
rs1023625	CAGAGCAAGACTCCATCTCA	ACT
rs1549297	GGTGACAGAGCAAGACTCC	CGT
rs3084894	CGAGTAGGTGGGACTACAG	ACT
rs8048228	TGAGCCGAGATGGCAACAC	ACG
rs7405432	CTACAGGCTAGGAGACAGAG	CGT
rs8054249	AAAGTGCTGGGATTACAGGC	ACT
rs8061047	CCTCCTGAGGAGCTGGTCT	ACT
rs7187220	GGCCCTTCCCCTGCACC	ACG
rs8046978	AGAGTTCAGCCGCCCGG	ACG
rs2288003	GTGACAAGACGTTCTGTGGC	ACT
rs7196421	CTCAGCCTCCCGAGTAGC	ACG
rs7196431	CGGGTTCACGCCATTCTCC	ACG
rs7203158	CAACCATGAGCGGTGAGGG	ACG
rs2303223	TTGAGCTCCGCATAGGCTTT	ACT
rs2032917	TGGAAATGTCTTGGTACAGGACA	ACT
rs8044134	CCTACACGTCCCCCCCC	ACT
rs4889531	CAACTCTGTCAGGTAAGTACT	ACT
rs4889532	CAAGAATCCCTATCTCAGAAAG	ACT
rs4889533	GGACAAGAATCCCTATCTCAG	ACT
rs881929	CTGCCTCTTGCCAGCTCTG	ACT
rs8047104	CAGAGACCTAGCCTACCTG	ACT
rs8047803	TTACAGGCGAGAGCCACCA	CGT
rs4644874	GGGCTCAAGTGATCCTCCC	CGT
rs2359673	ACTGCGACCTCTGCCTCC	ACG
rs4435271	GCTTCAGATGCTCCTCCACT	CGT
rs7197717	GCATGAGCCGTGACCAGC	CGT
rs2359674	GAATGTTTGTGTTCCCTGTCC	ACT
rs6565217	CCAGGGCCATACCCTTATGA	ACG
rs2303222	AAAGTGTACACCAAAGTAC	ACG
rs4889615	GCGGCGTCTTTGCACGCTA	CGT
rs4624197	AGAGAGTCTTTTGGGGTTTTTT	ACT
rs3751853	CCTACAGGTATAGCTAAGGAA	ACT
rs749671	ATTGTACCGCCGCTCCTC	ACG
rs749670	GGTGGTGGGCATCCCTTTC	ACG
rs3751855	AGAGCCCAGGCTGGAGAC	ACG
rs3751856	CCGTCTCGTGGCTGCGC	ACG
rs7196726	GTTAGGAACTGAGGAACCCAG	ACG
rs889550	AGCACTGCTGGAAGCCGC	ACT

dbSNP rs#	Extend Primer	Term Mix
rs750952	GCTGGCCTCTCCACCTCC	ACG
rs2077633	CCATATCTTCTCCTCTCCCC	ACG
rs7199949	CAGGCAGGTGGTGGTCAG	ACT
rs2032916	CCAAAGTTCCAGAGAGGTTAA	ACT
rs4468641	ATACGGGAGGCAGGCCCA	ACT
rs4889535	CAGACACCCTCTGATTGCAG	ACT
rs4316775	GAGGATCGCTTGAGCCCAA	ACT
rs4313819	GCTAGTCTTGAACCTTTGGG	ACT
rs6565218	CTCACGCTTGTAATCCCAGC	CGT
rs4318224	TTCCCTTGCAACCTGAGTTTT	ACG
rs1046030	GCCCAGGGAGGGAAGGTT	ACG
rs7294	TTGGTCCATTGTGTCATGTG	ACG
rs7200749	GAAGAACAGGATCCAGGCCA	ACT
rs2359612	CCATGTGTCAGCCAGGACC	ACT
rs8050894	CCAGCTAGCTGCTCATCAC	ACT
rs2884737	TCGCCAACACCCCCCTTC	CGT
rs1895514	CCCCTCTCGGGTTCAAGC	CGT
rs8060209	TGCCCAACATGGCGAAACC	ACG
rs8060217	CCATGCCCGGCTAATTTTTGT	ACT
rs7196161	AACTCCTGACCTCATGATCC	ACT
rs8062336	TCACTTCCTAGATGGGAAGG	ACG
rs8043778	CGCTCCTCACCTCCCAGA	ACG
rs2032915	TTCTTTCCCAACGTCCTGGA	ACT
rs4889616	GAACTGGAAGGAGGACAAGA	ACT
rs1045564	GTCCCTGAAGTCGGAGAAG	CGT
rs2303221	CTCTCCCTCCCGCCTACAT	ACG
rs1549296	TGCACGGGGCAGCCCCT	ACT
rs889555	AGCACCCCGGTTCTGTCC	ACT
rs5816521	CCAGTAGGTACAAGCACCC	ACT
rs749767	GACCAAGGATTTGGGCAAAG	ACT
rs2884738	CCAGGAGGTGGAGGTTGCA	ACT
rs2052581	GGATATGGAGGGCCGATTGT	ACT
rs4889617	GAGACTCCATCTCAAAAAAAAAA	ACT
rs4889619	GCAGAGGAATCGCTTGAACC	ACG
rs1978487	GTTCTGCAACATTTTTTCTTA	ACG
rs1978486	GGGTACAGATTGAAAAGTG	ACT
rs1978485	TTTTCTGCCTCAGCCTCC	ACG
rs4889620	ACCCCATCTGTACCAAAAAGA	ACG
rs4889621	CTGTGAGGTGGATCAGGTTG	ACT
rs3214477	GCAGAATCTGTGATGGAAAAAG	ACT
rs4527034	CCAGGGCAGCCAACTCCC	ACG
rs1060506	AAGTCTCCAGACACCCAGA	ACG
rs7200125	AGGCTCCATTTCAAAAAAAAAA	ACT
rs6565219	AAAGTGCTGGGATTACAGGC	ACT

dbSNP rs#	Extend Primer	Term Mix
rs889548	AGTCCTAATACAGTGGATGTC	ACT

Genetic Analysis

[0229] Allelotyping results from the discovery cohort are shown for cases and controls in Table 13. The allele frequency for the A2 allele is noted in the fifth and sixth columns for osteoarthritis case pools and control pools, respectively, where “AF” is allele frequency. The allele frequency for the A1 allele can be easily calculated by subtracting the A2 allele frequency from 1 (A1 AF = 1-A2 AF). For example, the SNP rs732173 has the following case and control allele frequencies: case A1 (G) = 0.55; case A2 (T) = 0.45; control A1 (G) = 0.58; and control A2 (T) = 0.42, where the nucleotide is provided in paranthesis. Some SNPs are labeled “untyped” because of failed assays.

TABLE 13

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs7500176	247	31077197	A/G			
rs6565212	1535	31078485	C/T			
rs8054046	2386	31079336	C/T			
rs8056842	6440	31083390	C/T			
rs732173	9133	31086083	G/T	0.45	0.42	0.382
rs732172	9143	31086093	A/G			
rs7188557	9471	31086421	A/T			
rs2288004	13150	31090100	C/G	0.52	0.45	0.026
rs4337310	13717	31090667	C/T	0.18	untyped	
rs2016554	14466	31091416	A/G			
rs6565213	15769	31092719	A/C			
rs7204762	16870	31093820	A/G			
rs4889529	18545	31095495	C/T			
rs6565214	18749	31095699	C/T			
rs7499674	19123	31096073	G/T			
rs6565215	20736	31097686	A/G			
rs1023623	21038	31097988	C/T	0.02	untyped	
rs1023624	21046	31097996	C/T	0.16	0.11	0.035
rs1023625	21050	31098000	C/T	0.32	NA	
rs1549297	21056	31098006	A/T			
rs3084894	21706	31098656	-/ACC			
rs8048228	23170	31100120	A/G	0.54	0.61	0.040
rs7405432	25028	31101978	A/T	0.35	0.43	0.025
rs8054249	27871	31104821	A/G			
rs8061047	28070	31105020	C/T	0.21	0.21	0.903
rs7187220	31717	31108667	A/G			
rs8046978	32019	31108969	A/G	0.34	0.28	0.083
rs2288003	32318	31109268	A/G			
rs7196421	33080	31110030	A/G			
rs7196431	33101	31110051	A/G			
rs7203158	34236	31111186	A/G			
rs2303223	34285	31111235	C/T	0.52	0.45	0.060
rs2032917	34818	31111768	C/G			

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs8044134	35168	31112118	C/G	0.97	0.97	0.856
rs4889531	37981	31114931	C/T			
rs4889532	38113	31115063	C/G			
rs4889533	38117	31115067	C/T			
rs881929	38481	31115431	G/T	0.38	0.34	0.228
rs8047104	38615	31115565	C/G	0.60	0.65	0.117
rs8047803	38944	31115894	A/C	0.35	0.33	0.437
rs4644874	39288	31116238	A/C			
rs2359673	41385	31118335	C/T	0.18	0.20	0.563
rs4435271	42136	31119086	A/T			
rs7197717	42185	31119135	A/C			
rs2359674	42353	31119303	A/G	0.22	0.18	0.122
rs6565217	42434	31119384	A/G	0.35	0.33	0.608
rs2303222	44580	31121530	A/G	0.60	0.52	0.022
rs4889615	44675	31121625	A/T			
rs4624197	45739	31122689	G/T			
rs3751853	46439	31123389	C/T			
rs749671	47457	31124407	C/T	0.32	0.37	0.095
rs749670	47735	31124685	C/T			
rs3751855	50319	31127269	C/T	0.53	0.57	0.287
rs3751856	50708	31127658	A/G			
rs7196726	51185	31128135	A/G	0.41	0.37	0.258
rs889550	53002	31129952	A/G			
rs750952	53064	31130014	C/T	0.43	0.41	0.535
rs2077633	53637	31130587	A/G			
rs7199949	55274	31132224	C/G	0.46	0.53	0.051
rs2032916	55825	31132775	C/T			
rs4468641	55986	31132936	A/C	0.26	0.25	0.902
rs4889535	56684	31133634	C/G			
rs4316775	57653	31134603	C/T			
rs4313819	57659	31134609	C/G			
rs6565218	57692	31134642	G/T			
rs4318224	57775	31134725	C/T			
rs1046030	61313	31138263	C/T			
rs7294	61431	31138381	A/G	0.38	0.37	0.669
rs7200749	61699	31138649	A/G			
rs2359612	62906	31139856	A/G	0.56	0.48	0.017
rs8050894	63619	31140569	C/G	0.48	0.45	0.320
rs2884737	64664	31141614	A/C	0.68	0.60	0.016
rs1895514	68452	31145402	G/T			
rs8060209	69665	31146615	C/T			
rs8060217	69681	31146631	C/T			
rs7196161	70091	31147041	A/G			
rs8062336	74637	31151587	A/G			
rs8043778	74760	31151710	A/G			
rs2032915	76523	31153473	A/G	0.43	0.41	0.505
rs4889616	78559	31155509	C/G			
rs1045564	79549	31156499	A/C			
rs2303221	79882	31156832	C/T			
rs1549296	81339	31158289	A/G			
rs889555	81681	31158631	C/T	0.49	0.50	0.740
rs5816521	81696	31158646	-/G			
rs749767	83517	31160467	C/T	0.28	0.36	0.020
rs2884738	85431	31162381	A/C			
rs2052581	86332	31163282	C/T			
rs4889617	87358	31164308	A/G			
rs4889619	87725	31164675	C/T			

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs1978487	89052	31166002	A/G	0.62	0.57	0.124
rs1978486	90020	31166970	A/G			
rs1978485	90231	31167181	A/G	0.90	0.88	0.513
rs4889620	90284	31167234	A/G			
rs4889621	90447	31167397	C/T			
rs3214477	90601	31167551	-/G			
rs4527034	90724	31167674	A/G	0.37	0.43	0.079
rs1060506	92559	31169509	C/T	0.29	0.28	0.720
rs7200125	95176	31172126	A/G			
rs6565219	95195	31172145	C/T			
rs889548	96822	31173772	A/G	0.54	0.51	0.320
rs6145813	Not mapped	Not mapped	/TTTTT TTTTT	0.33	0.32	0.909

[0230] Allelotyping results were considered particularly significant with a calculated p-value of less than or equal to 0.05 for allelotyping results. These values are indicated in bold. The allelotyping p-values were plotted in Figure 1A for the discovery cohort. The position of each SNP on the chromosome is presented on the x-axis. The y-axis gives the negative logarithm (base 10) of the p-value comparing the estimated allele in the case group to that of the control group. The minor allele frequency of the control group for each SNP designated by an X or other symbol on the graphs in Figure 1C can be determined by consulting Table 13. For example, the left-most X on the left graph is at position 31077197. By proceeding down the Table from top to bottom and across the graphs from left to right the allele frequency associated with each symbol shown can be determined.

[0231] To aid the interpretation, multiple lines have been added to the graph. The broken horizontal lines are drawn at two common significance levels, 0.05 and 0.01. The vertical broken lines are drawn every 20kb to assist in the interpretation of distances between SNPs. Two other lines are drawn to expose linear trends in the association of SNPs to the disease. The light gray line (or generally bottom-most curve) is a nonlinear smoother through the data points on the graph using a local polynomial regression method (W.S. Cleveland, E. Grosse and W.M. Shyu (1992) Local regression models. Chapter 8 of Statistical Models in S eds J.M. Chambers and T.J. Hastie, Wadsworth & Brooks/Cole.). The black line provides a local test for excess statistical significance to identify regions of association. This was created by use of a 10kb sliding window with 1kb step sizes. Within each window, a chi-square goodness of fit test was applied to compare the proportion of SNPs that were significant at a test wise level of 0.01, to the proportion that would be expected by chance alone (0.05 for the methods used here). Resulting p-values that were less than 10^{-8} were truncated at that value.

[0232] Finally, the exons and introns of the genes in the covered region are plotted below each graph at the appropriate chromosomal positions. The gene boundary is indicated by the broken

horizontal line. The exon positions are shown as thick, unbroken bars. An arrow is placed at the 3' end of each gene to show the direction of transcription.

Example 5

Chromosome 4 Region Proximal SNPs

[0233] SNP rs1957723 is associated with osteoarthritis and is described in Table A. One hundred-thirty additional allelic variants proximal to rs1957723 were identified and subsequently allelotyped in osteoarthritis case and control sample sets as described in Examples 1 and 2. The polymorphic variants are set forth in Table 14. The chromosome positions provided in column four of Table 14 are based on Genome "Build 34" of NCBI's GenBank.

TABLE 14

dbSNP rs#	Chromosome	Position in SEQ ID NO: 2	Chromosome Position	Allele Variants
rs3849023	4	211	36870611	g/t
rs1444311	4	7217	36877617	a/g
rs2044295	4	7895	36878295	a/c
rs2166093	4	13308	36883708	c/t
rs2376334	4	14279	36884679	g/t
rs1444320	4	17026	36887426	c/t
rs2044294	4	18271	36888671	a/g
rs1899864	4	20417	36890817	c/t
rs1562094	4	21843	36892243	a/g
rs1562098	4	22069	36892469	a/g
rs1562097	4	22145	36892545	a/g
rs1562096	4	22519	36892919	a/g
rs1562095	4	22539	36892939	a/g
rs1444319	4	23236	36893636	a/c
rs1444318	4	23256	36893656	a/c
rs1025938	4	23402	36893802	c/t
rs1025937	4	23499	36893899	a/c
rs1025936	4	23620	36894020	c/t
rs1020333	4	23871	36894271	a/t
rs2120654	4	24136	36894536	c/g
rs2588547	4	25427	36895827	a/g
rs2044293	4	25866	36896266	g/t
rs2760324	4	26541	36896941	a/g
rs2588546	4	26576	36896976	g/t
rs2588545	4	26689	36897089	a/g
rs2760328	4	26720	36897120	a/c
rs2588544	4	27113	36897513	c/t
rs2760331	4	27164	36897564	c/t
rs2588543	4	27186	36897586	a/g
rs2588542	4	28341	36898741	a/t

dbSNP rs#	Chromosome	Position in SEQ ID NO: 2	Chromosome Position	Allele Variants
rs2588541	4	29160	36899560	c/t
rs2588540	4	29844	36900244	a/g
rs2760336	4	30665	36901065	g/t
rs2760337	4	30830	36901230	a/g
rs2028732	4	31061	36901461	a/c
rs2588538	4	31523	36901923	c/t
rs1992617	4	32326	36902726	c/t
rs1998469	4	32346	36902746	a/g
rs1998470	4	32358	36902758	c/t
rs1975498	4	34909	36905309	c/t
rs1562093	4	34975	36905375	a/g
rs1975497	4	35066	36905466	c/t
rs1562092	4	35096	36905496	g/t
rs2248788	4	35375	36905775	c/t
rs1899862	4	36304	36906704	a/g
rs2588532	4	36712	36907112	a/t
rs1885878	4	36770	36907170	c/t
rs986648	4	37342	36907742	c/t
rs986647	4	37412	36907812	c/t
rs1010010	4	37884	36908284	a/g
rs1010009	4	38077	36908477	a/c
rs2760325	4	38300	36908700	c/t
rs2588531	4	38301	36908701	c/t
rs1838388	4	41189	36911589	c/t
rs1975495	4	44408	36914808	c/t
rs2181491	4	44493	36914893	a/c
rs1975496	4	44571	36914971	a/g
rs2181492	4	44670	36915070	a/g
rs2224719	4	45219	36915619	a/g
rs2224720	4	45258	36915658	c/t
rs1951770	4	47261	36917661	a/g
rs2296040	4	48473	36918873	a/c
rs1957723	4	48771	36919171	a/g
rs1957725	4	55292	36925692	c/t
rs2889346	4	56479	36926879	a/g
rs1885879	4	56747	36927147	a/c
rs1957726	4	60620	36931020	g/t
rs1957727	4	60688	36931088	a/c
rs1885880	4	61058	36931458	a/c
rs1885881	4	61129	36931529	c/t
rs942108	4	61577	36931977	c/t
rs1951771	4	61961	36932361	a/g
rs2376323	4	63351	36933751	g/t
rs2013358	4	63926	36934326	a/g
rs2181494	4	65798	36936198	a/g
rs1957728	4	66043	36936443	a/c
rs1957729	4	66044	36936444	a/g

dbSNP rs#	Chromosome	Position in SEQ ID NO: 2	Chromosome Position	Allele Variants
rs1957730	4	66246	36936646	c/t
rs1957731	4	66318	36936718	c/t
rs1998468	4	66547	36936947	g/t
rs1957732	4	71238	36941638	c/t
rs1957733	4	71283	36941683	a/g
rs2376322	4	71492	36941892	a/g
rs2889345	4	72274	36942674	a/g
rs1815267	4	73762	36944162	a/t
rs1957734	4	74209	36944609	g/t
rs1957735	4	75284	36945684	a/t
rs1957736	4	77347	36947747	a/c
rs1957737	4	77589	36947989	c/t
rs1957738	4	78096	36948496	a/g
rs1957739	4	78606	36949006	a/g
rs1957740	4	78862	36949262	g/t
rs1957741	4	79135	36949535	a/g
rs1957742	4	79146	36949546	a/g
rs1957743	4	79456	36949856	c/t
rs1957744	4	79609	36950009	a/g
rs1957745	4	80086	36950486	a/g
rs1957746	4	80119	36950519	a/g
rs1957747	4	80766	36951166	c/t
rs2146670	4	81110	36951510	a/g
rs2146671	4	81269	36951669	a/t
rs1957748	4	81668	36952068	c/t
rs2162307	4	82433	36952833	c/t
rs1962839	4	82559	36952959	c/g
rs2376315	4	83298	36953698	c/t
rs1426410	4	83821	36954221	a/g
rs1895921	4	84121	36954521	c/t
rs1895922	4	84147	36954547	c/t
rs1035779	4	84543	36954943	a/g
rs1035780	4	84554	36954954	a/g
rs1035781	4	84691	36955091	a/g
rs1035782	4	84727	36955127	a/g
rs1426411	4	85678	36956078	c/t
rs1834602	4	86699	36957099	c/t
rs1834603	4	86700	36957100	a/g
rs1834604	4	86792	36957192	a/g
rs1834605	4	86832	36957232	a/g
rs2162308	4	87045	36957445	a/g
rs1365341	4	87140	36957540	a/g
rs1820458	4	87365	36957765	a/c
rs1469310	4	88342	36958742	c/t
rs3057879	4	88498	36958898	-/tca
rs1469311	4	88589	36958989	a/g
rs768326	4	95502	36965902	a/g

dbSNP rs#	Chromosome	Position in SEQ ID NO: 2	Chromosome Position	Allele Variants
rs1863523	4	96968	36967368	c/t
rs1469312	4	97448	36967848	c/t
rs1469313	4	97568	36967968	c/t
rs1951773	4	98724	36969124	c/t
rs2120655	4	Not mapped	Not mapped	t/g
rs2181495	4	Not mapped	Not mapped	g/a

Assay for Verifying and Allelotyping SNPs

[0234] The methods used to verify and allelotype the 130 proximal SNPs of Table 14 are the same methods described in Examples 1 and 2 herein. The primers and probes used in these assays are provided in Table 15 and Table 16, respectively.

TABLE 15

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs3849023	ACGTTGGATGGGTAATTGCTAACCATGTTC	ACGTTGGATGGACCCAGTCAAGTCAATAAAC
rs1444311	ACGTTGGATGGCCTATTGGTTTAACTAGGC	ACGTTGGATGTCTGGCTTCTTCAGGAGTTC
rs2044295	ACGTTGGATGCCACACCACTACTATTCAAG	ACGTTGGATGGTGGTGTGTTAGAAGGTTAC
rs2166093	ACGTTGGATGAAAATCCTGGAGATGGATGG	ACGTTGGATGTAGGTGTACAGTTCAGTGTCT
rs2376334	ACGTTGGATGTCTCAGAGAACCAGCTTTTG	ACGTTGGATGGGGAATATTAAACATTGGGG
rs1444320	ACGTTGGATGTAATTCTCTCCTCCAAATGC	ACGTTGGATGCTAGAAACAAAAGACTACATG
rs2044294	ACGTTGGATGAACCTAAATCTCCTCAAGCC	ACGTTGGATGTTCTGACCACTTCTCTATGG
rs1899864	ACGTTGGATGTTTATAGGCGTGGGCAATCG	ACGTTGGATGTTGTCAGAAAGTGTCTGCGGCC
rs1562094	ACGTTGGATGTGGATTCTTTCTTGAAGAC	ACGTTGGATGGCAACAAGAACTTAATGC
rs1562098	ACGTTGGATGTCTGAGTCCGAGTGATCATC	ACGTTGGATGAAACAATTAGCAGGGCACAG
rs1562097	ACGTTGGATGCACAGGATCTTACTCTGTTG	ACGTTGGATGCGGACTCAGAAATTCAAGTC
rs1562096	ACGTTGGATGACCCAGGGCATGTTATATAG	ACGTTGGATGTTTCTCTCTGGTACCCTCTC
rs1562095	ACGTTGGATGTGTTAGTAACCCAGGGCATG	ACGTTGGATGTGACAGATGCCACCAGTTAC
rs1444319	ACGTTGGATGTTCAACTTTAGCCTCTGGGC	ACGTTGGATGCCCTGCAAAGTCAAAGGAAC
rs1444318	ACGTTGGATGCTCTGGGCAATTATCAAGCC	ACGTTGGATGAGTTCGCTGATGTGTTTGGG
rs1025938	ACGTTGGATGCAGGTAAGAAAAGCTTTTTGG	ACGTTGGATGCCCTGCTAATGACTGAATTTT
rs1025937	ACGTTGGATGGAATAGGAAAGGTAGTATACC	ACGTTGGATGAAATTCAGTCATTAGCAGGG
rs1025936	ACGTTGGATGTCTCCAGGTAGATGAGTCAG	ACGTTGGATGCCACACACCAAAGCAATCAC
rs1020333	ACGTTGGATGGCATCTCTTCAATCTGGACG	ACGTTGGATGGTGGATCACAGAAGTCAGAG
rs2120654	ACGTTGGATGACCAGAAAGACCAGGGCATG	ACGTTGGATGAACCTTTAGCTCTTCTCCCC
rs2588547	ACGTTGGATGTCACAAATGTAATATAAATC	ACGTTGGATGGATAGCTACGTTTAAAAATG
rs2044293	ACGTTGGATGTGTCAACAATACAAGACTAA	ACGTTGGATGTGCACTGGACTTTTTTTTTT
rs2760324	ACGTTGGATGACAAACCACTGGTTGAGGAG	ACGTTGGATGCCTCACGAATCCAACAGAAC
rs2588546	ACGTTGGATGCTTAGAGGATGGAGTCAGTC	ACGTTGGATGTACTACCAGAGATGCTGGTG
rs2588545	ACGTTGGATGCAACACAGCTACAGTGCATC	ACGTTGGATGTGGGTAAAGGGAAAAGAAGG
rs2760328	ACGTTGGATGGCCATAAAATTGGGTAAAGGG	ACGTTGGATGGCATCTATTTGACACCAACG
rs2588544	ACGTTGGATGTAAGAATTAGCATGTGAAAG	ACGTTGGATGTTTGTGCACAAAGAATTTGG
rs2760331	ACGTTGGATGAAACAGTATGCCTTTTGTGC	ACGTTGGATGCTTCTCGTAATTTTACATGAC

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs2588543	ACGTTGGATGGTGCCAAATTCTTTGTGCAC	ACGTTGGATGCTAAGATAGGTAGATACCAG
rs2588542	ACGTTGGATGTGGCAGCAAAGCTTAAGCTC	ACGTTGGATGTCCACAGTCACCTCTCATT
rs2588541	ACGTTGGATGTGACAAGGTCTATGTCAGGG	ACGTTGGATGGGCATTGTTCATGGTGATGAG
rs2588540	ACGTTGGATGTGCTGTATGATCCAGCAATC	ACGTTGGATGGGTGCAAATACTGTCTCTTC
rs2760336	ACGTTGGATGAAGCTGAGGCAGGAGAATGG	ACGTTGGATGTGTTTTGAGACGGAGTCTCG
rs2760337	ACGTTGGATGGGTGTTTCAACTGATACAAG	ACGTTGGATGACTACCATTCTACTCTCTGC
rs2028732	ACGTTGGATGTTCTTGACAGCTAAATAGG	ACGTTGGATGGCCATTGTCTGTTTTCTTGTT
rs2588538	ACGTTGGATGTATCTTCTGGGAAGCCTTTC	ACGTTGGATGGACTTGAAATCACTCCATGC
rs1992617	ACGTTGGATGGGAGGACATTGCCTTCAAAG	ACGTTGGATGCTGACCTTCTGTCTAGTCAC
rs1998469	ACGTTGGATGTATATGCCAAGGACCAACGG	ACGTTGGATGCTGACCTTCTGTCTAGTCAC
rs1998470	ACGTTGGATGATTTCCCCCATTAAGCTTTG	ACGTTGGATGGAAAAGTATTATATGCCAAGG
rs1975498	ACGTTGGATGAGCTCTCTTTTTGCCTGCTG	ACGTTGGATGAGGAGGCTTCACAATCATGG
rs1562093	ACGTTGGATGTGATTGTGAAGCCTCCTCTG	ACGTTGGATGAAAGACATACCCAAGACTGG
rs1975497	ACGTTGGATGTCAGCAGCATGAAAAGTAC	ACGTTGGATGCATTAGACTTTTTCTGGGG
rs1562092	ACGTTGGATGTTCCAGTGACTGGACCATAG	ACGTTGGATGTCAGCAGCATGAAAAGTAC
rs2248788	ACGTTGGATGGGGAAAAGAAAAAGACTTCC	ACGTTGGATGGTAGTAGCTGCTTCTAAAAG
rs1899862	ACGTTGGATGTAATCTCCATAATAAGTGC	ACGTTGGATGGCTACAAAAGAAAATGAATAC
rs2588532	ACGTTGGATGCAAACAATAGTGGCTGAGAG	ACGTTGGATGTTGTAGCACAGGCGCATAG
rs1885878	ACGTTGGATGTGACTCAGCGAGTTGTAGC	ACGTTGGATGAGCCAGATTGGGTGCTTTTC
rs986648	ACGTTGGATGTGAGAAAGCTTTCTGAGGAC	ACGTTGGATGGGTTTTCTGTTGTGAATGGG
rs986647	ACGTTGGATGACACACTCTTCTCAAGCAG	ACGTTGGATGCTTATTGTCTCAGAAAGC
rs1010010	ACGTTGGATGACTGTAGCTAAGTTGGCAT	ACGTTGGATGTTCAACCAACCAATAAGGC
rs1010009	ACGTTGGATGTCTCATCAGCTCTTTCCTGG	ACGTTGGATGAAAGGGATGAGGAAGTGAGG
rs2760325	ACGTTGGATGATCCCCAGCATGTAGCATAG	ACGTTGGATGCTGCCCATAGTCTCTTCTG
rs2588531	ACGTTGGATGATCCCCAGCATGTAGCATAG	ACGTTGGATGCTGCCCATAGTCTCTTCTG
rs1838388	ACGTTGGATGGTACCTCATGGATATTACAC	ACGTTGGATGTTGGTGTGTTATAAATGAC
rs1975495	ACGTTGGATGCAGGTCAGGAGTTAAGACC	ACGTTGGATGAGCTGGGATTACAGTCATGC
rs2181491	ACGTTGGATGGTACCTAATATATGCTTCTGG	ACGTTGGATGTTATTCCTGCTTACTTTCC
rs1975496	ACGTTGGATGTATATTAGGTACAGTGTGGC	ACGTTGGATGCAACCAACTTCACTGAAAGC
rs2181492	ACGTTGGATGCTTGCAAGGAAGAGGAAGAAG	ACGTTGGATGACAATCACCTTTGGAGGCAG
rs2224719	ACGTTGGATGTCAAGGGTGTAGATGTGTAG	ACGTTGGATGCCAGAGAGGAGTAATGGTAT
rs2224720	ACGTTGGATGCCAATTACTCAAGGGTGTAG	ACGTTGGATGAATTCAGTACAGACAGAGGG
rs1951770	ACGTTGGATGCCTGGGAACCTTCAGCTTTTC	ACGTTGGATGTGGCACAGCAGGAATATCAG
rs2296040	ACGTTGGATGGGGCATCATGAAATGCAGAC	ACGTTGGATGGCATGTACAGGAAAGCAGTG
rs1957723	ACGTTGGATGTACTCACTTGTGTACTGCTC	ACGTTGGATGGCTGCAGCGTCACATTAATC
rs1957725	ACGTTGGATGTTATTGGAATCTCCAGGTC	ACGTTGGATGAAGATGATTAGTCCAGCCTG
rs2889346	ACGTTGGATGTGACTGACTTCCTAGGTCAG	ACGTTGGATGTGACAGTGTGAGTGGCAG
rs1885879	ACGTTGGATGTTACCCCTTCACATCTGAT	ACGTTGGATGCTACAAGGAAGATAACAGAG
rs1957726	ACGTTGGATGAAATTCAGCCACTCAACCAG	ACGTTGGATGAAGTGGTTGGGATTTGTGAG
rs1957727	ACGTTGGATGGCCAACGTATCTTTAAACCC	ACGTTGGATGGTTTTGTCTTGGTTCTCATC
rs1885880	ACGTTGGATGTGGAATGCCCAAGATTTCA	ACGTTGGATGCTGGAATCCCAAGGTTCTCTG
rs1885881	ACGTTGGATGTAGACGTGTTCTGCATCATG	ACGTTGGATGATGAAATCTTGGGGCATTCC
rs942108	ACGTTGGATGGAGCTGTTAGGGTAGAAATG	ACGTTGGATGGTCTTGGACTAATTTTGACC
rs1951771	ACGTTGGATGGGCATTCCCTTTTGTCTAAG	ACGTTGGATGAGTAAACAAGGACTAGAGCC
rs2376323	ACGTTGGATGTCCTTACTTGCTAGCACTGC	ACGTTGGATGGCATCCCTTGGTGACTGATA
rs2013358	ACGTTGGATGGGAATTTTAGGAGTACTGTAG	ACGTTGGATGGCCAACCATAGAACCTAAATC
rs2181494	ACGTTGGATGATTCAATTACCTCCCACTGG	ACGTTGGATGTATCCCCACCAATGTAC

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs1957728	ACGTTGGATGAAATAGATCCCAACCAAGGG	ACGTTGGATGGTAACATTTACCTAAGCGGG
rs1957729	ACGTTGGATGAAATAGATCCCAACCAAGGG	ACGTTGGATGGTAACATTTACCTAAGCGGG
rs1957730	ACGTTGGATGGGTCTAAACATGAGAGACTC	ACGTTGGATGTCTTTATGGATATAGGGTCC
rs1957731	ACGTTGGATGTATTGGAACCTGGTACCTGG	ACGTTGGATGGACCTGAATCATGTCTCCAG
rs1998468	ACGTTGGATGTATAAAGCCTCAAAAGTGGG	ACGTTGGATGACCTTATTCCAGAATGAAAC
rs1957732	ACGTTGGATGAAGAGAGGAGTTTATTGGCC	ACGTTGGATGCGGCCTGATCTTTATTTTCG
rs1957733	ACGTTGGATGCTATCAAGACTCTGATTGCC	ACGTTGGATGTGTTTGCAGGTAAACTTGGC
rs2376322	ACGTTGGATGTCGTTCTCTCTCTGTGCATG	ACGTTGGATGTTAGTCAGATGCTTGGTGAG
rs2889345	ACGTTGGATGTGGAATCCCAAACCTTTTACG	ACGTTGGATGTTCTTGCTAAATGTAGGCC
rs1815267	ACGTTGGATGCAGGAAAGGGCTACTATCAG	ACGTTGGATGGTAGGCCAAACTAGCTTTGG
rs1957734	ACGTTGGATGCTACCCCTGCCTTATAATTC	ACGTTGGATGCAAGTGGTAAAAGGATGTGG
rs1957735	ACGTTGGATGAGCTTCCCATGGTTATAGAG	ACGTTGGATGCTGAAAACAATACCGGTCTC
rs1957736	ACGTTGGATGCTGAAGCAAAGATTTCTCTC	ACGTTGGATGAGCATCTTTTGCTGTCACTG
rs1957737	ACGTTGGATGACATGGAAGCTGAAGCCAAG	ACGTTGGATGCAGAGCTTTGACCTTACTCC
rs1957738	ACGTTGGATGATGTCCCTTAAAAGGCTGCC	ACGTTGGATGCAGATGATCTTGCTTCCCAG
rs1957739	ACGTTGGATGTCACTGCCTGAGTGCTTTAG	ACGTTGGATGCTGATGGCCTGAGAACTAAG
rs1957740	ACGTTGGATGGCCCAGTCAAGTTGACATAC	ACGTTGGATGCACCTGCTCCAGTTATATAC
rs1957741	ACGTTGGATGAGGAGCATTATCCCTATTAG	ACGTTGGATGCCTCTTAGTAAAATATGGATG
rs1957742	ACGTTGGATGGGATGATATCTACTTTGTACG	ACGTTGGATGGACTCCATCTGAGATGTTAG
rs1957743	ACGTTGGATGCAACTGTCTTGATTTGAAG	ACGTTGGATGGACAGACTTTTCATTGTTTTC
rs1957744	ACGTTGGATGTCAGTGTACCCTGTAATGCC	ACGTTGGATGTGCCAGCAGTGAGTAATTG
rs1957745	ACGTTGGATGGTTTAGAAAGTGTGGGTTCC	ACGTTGGATGAGCAAATGCAGCTTATTACC
rs1957746	ACGTTGGATGTCTCATACAACATAGTTAGC	ACGTTGGATGGGTTTAGGTTTGGTTTGATG
rs1957747	ACGTTGGATGGTCACTCAAGATAACAGTTCC	ACGTTGGATGTTACCTAACGTGAAGGTAGC
rs2146670	ACGTTGGATGCCTAACACATCTTTATGAGC	ACGTTGGATGCTCATAAGATATGCTAAGCAC
rs2146671	ACGTTGGATGATGAGGAGCAACTAGAAGGC	ACGTTGGATGAAAGGGCTGGAAGAAACAGG
rs1957748	ACGTTGGATGTGAAGTTTGTAGTAGGGAGC	ACGTTGGATGTTCTGTACACAAACACTCC
rs2162307	ACGTTGGATGACATGCGGTGCCTGGCCCTTT	ACGTTGGATGCCTTTGTAGGGACATGGATG
rs1962839	ACGTTGGATGGGCTGCATAGTATTCCATGG	ACGTTGGATGAGGGAATCCTTTCCCCATTG
rs2376315	ACGTTGGATGTGGCCTTGGATTTCTTCCAC	ACGTTGGATGAGAATTGGACAGAGTGGCAG
rs1426410	ACGTTGGATGGAGAAAGTTGCATCTTGCCC	ACGTTGGATGGGGAAGTTTTACCTTGGCTC
rs1895921	ACGTTGGATGGGTGATGGTGTTTGAGGTAC	ACGTTGGATGATTAGGCTTCTCCCACCATC
rs1895922	ACGTTGGATGCAATGCATTAGGCTTCTCCC	ACGTTGGATGGAGGTACATTTCTCAGGCAG
rs1035779	ACGTTGGATGGAGAATCACTTGAACCCGGG	ACGTTGGATGTGGAGTGCAGTGGCATGATC
rs1035780	ACGTTGGATGTTTGAGATGGAGTCTCGCTC	ACGTTGGATGAATCACTTGAACCCGGGAGG
rs1035781	ACGTTGGATGGGAAGATGCTGACTCTGAAC	ACGTTGGATGCCTTGACTGTTTAGGGATCC
rs1035782	ACGTTGGATGGGATCCCTAAACAGTCAAGG	ACGTTGGATGAGTTGGCTAGACTTGCCTTC
rs1426411	ACGTTGGATGCAAGAGTGCTACACAAGTCG	ACGTTGGATGTGTACCTTGGTCAGGTGATC
rs1834602	ACGTTGGATGGATGGGCCCTATTTTCTTG	ACGTTGGATGCTTTTCCAACCCAGTAATGTC
rs1834603	ACGTTGGATGGATGGGCCCTATTTTCTTG	ACGTTGGATGTCTTTTCCAACCCAGTAATG
rs1834604	ACGTTGGATGGAAAGACATTACTGGGTTGG	ACGTTGGATGAGAATTCTTCTGACTGTGG
rs1834605	ACGTTGGATGGCCACAGTCAGGAAGAATT	ACGTTGGATGTTGTGGAGACTGGCCAAAAG
rs2162308	ACGTTGGATGTAAAGAAACAGAGGGACACC	ACGTTGGATGTATGATCAGAGTCATCAGGG
rs1365341	ACGTTGGATGTCCCTCTGTTTCTTTAGGCA	ACGTTGGATGCATCTCCCTGGTAGCATT
rs1820458	ACGTTGGATGCACCCTCAGACTTGGAAATG	ACGTTGGATGGTCAGGTGACTCTATTACAGC
rs1469310	ACGTTGGATGTACTACAGCGTGTTAGCAG	ACGTTGGATGTGTCAAAGGGAGAGTTAGAG
rs3057879	ACGTTGGATGGGCACATTGGAATAAAGCC	ACGTTGGATGACGGCATGAACAATTCTCAG

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs1469311	ACGTTGGATGCCTGAGAATTGTTTCATGCCG	ACGTTGGATGTTTTCAAGTGTCTCTCCAGG
rs768326	ACGTTGGATGAATTAGCCAGGCATGGTGTG	ACGTTGGATGACATCCTAGGCTCAAGTGAC
rs1863523	ACGTTGGATGGGCAGACACATTCTATTCCG	ACGTTGGATGGGGAAAGGTGTGCTGAGTAA
rs1469312	ACGTTGGATGCATTTTCGTCAGCATTCTAGC	ACGTTGGATGGGACTCATGTCTCTCTTTGG
rs1469313	ACGTTGGATGAGTGAGGGAGAAAAGTGAAC	ACGTTGGATGCCTAACTTCTCTCCAATCTC
rs1951773	ACGTTGGATGAAGGTTCAAGTTACCGCATG	ACGTTGGATGCACTGTGGTCCATGAAAAA
rs2120655	ACGTTGGATGACAGGGTTTCTGCATGTTGC	ACGTTGGATGACGCCTGTAATCCAGCACT
rs2181495	ACGTTGGATGGAATTGTGGGAGTTACAATTC	ACGTTGGATGGAATCAAGCTAATTAACATGTG

TABLE 16

dbSNP rs#	Extend Primer	Term Mix
rs3849023	CTCATAACATAAGAAGTTGATGC	CGT
rs1444311	CTAGGCATGCTAGCTTGGC	ACT
rs2044295	CACTACTATTCAAGATTACCCTTT	ACT
rs2166093	GGTGGTGATGGCTGCACAA	ACG
rs2376334	TCAGAGAACCAGCTTTTGATTTCA	ACT
rs1444320	GCCTAGACCCCGTGCAAC	ACG
rs2044294	CTCCTCAAGCCAATAGGTCTTA	ACG
rs1899864	CGCACCTGGCCGAAAATAAC	ACT
rs1562094	AACCTGCAAAAGATTTACACTTGC	ACT
rs1562098	TCCTGCCTCAGCCTTCCTAGA	ACT
rs1562097	ACTCTGTTGTTTCAGGCTGGGGT	ACT
rs1562096	TAAGCTAGCTAGTAAGTGGTG	ACT
rs1562095	ATGTTATATAGAACATCCCTTTTT	ACT
rs1444319	TCTGGGCAATTATCAAGCCTTT	ACT
rs1444318	CTTTGCATTTTCCTGAGTTCCTTT	ACT
rs1025938	AAGAAAAGCTTTTTGTTTGGG	ACT
rs1025937	GGTAGTATACCTAAAAAACAGC	CGT
rs1025936	TCAAAGGACACCCAGCATTCA	ACG
rs1020333	ACGTTTATCTGTAACCTTTCCA	CGT
rs2120654	GAAAGACCAGGGCATGATTAGA	ACT
rs2588547	ACAAATGTAATATAAATCAAGCTC	ACG
rs2044293	ACCAGCCTGGGTAACATAGCCA	ACT
rs2760324	GGTTGAGGAGAAGCACCAGCA	ACG
rs2588546	TACAATTTCTAGCCTTAATAAGAT	ACT
rs2588545	TACAGTGCATCTATTTGACACCAA	ACG
rs2760328	AAATTGGGTAAAGGGAAAAGAAG	ACT
rs2588544	ATTAGCATGTGAAAGACTTCTC	ACT
rs2760331	AGTATGCCTTTTGTGCACAAAGA	ACT
rs2588543	ATTCTTTGTGCACAAAAGGCATA	ACG
rs2588542	GCTTAAGCTCTTACAGGCAG	CGT
rs2588541	AGGTCTATGTCAGGGAAAACCTTA	ACG

dbSNP rs#	Extend Primer	Term Mix
rs2588540	GATCCAGCAATCCCACTGAT	ACG
rs2760336	AGGCGGAGCTTGCAGTGAG	ACT
rs2760337	CACCAATACTGTATGATTCTTTT	ACT
rs2028732	CAGCTAAATAGGGCTTGAGTCAAT	CGT
rs2588538	AATTTGTACAAATTTATGGGGTAT	ACT
rs1992617	ATTGCCTTCAAAGAACATCAAAGC	ACG
rs1998469	GACCAACGGGAGGACATTG	ACG
rs1998470	CTTTGAAGGCAATGTCCTCC	ACG
rs1975498	TTTTGCCTGCTGCTATCCAC	ACT
rs1562093	CTCCTCTGCCATGTGGAAC	ACG
rs1975497	AAAACCTGACTAATACACACTGTT	ACT
rs1562092	TTTGGTTAATGGACATTTAGACT	ACT
rs2248788	TGTGGGATTTTATTATTTTCATCA	ACT
rs1899862	TAAGTGCATAACTTGTCTTTGAGG	ACT
rs2588532	ATAGTGGCTGAGAGCCAGAT	CGT
rs1885878	GCGAGTTTGTAGCACAGGC	ACT
rs986648	GTACATGTAATGCTAGTAAAGAAA	ACG
rs986647	CTCTTTCTCAAGCAGGAGTTA	ACG
rs1010010	AGCTAAGTTGGCATGTGGGA	ACT
rs1010009	CCTGGCTACCTTCCAAAAAG	ACT
rs2760325	TCTCAGGAAGTATGAAATAAATAG	ACG
rs2588531	CTCAGGAAGTATGAAATAAATAGT	ACT
rs1838388	TCATGGATATTTACACCTACTAC	ACT
rs1975495	AGGAGTTTAAGACCAGCCTG	ACT
rs2181491	TGCTTCTGGATTTTAAATGATCAC	ACT
rs1975496	ATGATCAAATCATTTTGAGGGC	ACT
rs2181492	GTTGCATTGCTATGGTCTGC	ACT
rs2224719	CATATATCCCTCTGTCTGTAC	ACG
rs2224720	GGGTGTAGATGTGTAGATTTATA	ACT
rs1951770	ACAAGCATTAGAGACTTGATTG	ACG
rs2296040	CTTTGTTTCTAAAATCTGATAGTC	ACT
rs1957723	AGCATGGCATAGGCACTGG	ACG
rs1957725	GCGAGGAAAGACCTGTTCTA	ACG
rs2889346	GGTCAGCTCAGCTGGTTTTT	ACG
rs1885879	CACATCTGATGCTCTCCTAAA	ACT
rs1957726	GCCACTCAACCAGTAGGAAA	ACT
rs1957727	GTATCTTTAAACCCTCACAAAT	CGT
rs1885880	TTACGTTAGTCTGCCTACTTCCA	ACT
rs1885881	TGGGCTATCAATGATGGAAAC	ACT
rs942108	AAATGAAATAGAATTGTGTACTTC	ACT
rs1951771	TCCCTTTTGTCTAAGAATATTAG	ACG
rs2376323	CTAGCACTGCCAAGTGCAAC	ACT
rs2013358	TTTTAGGAGTACTGTAGAACACA	ACG
rs2181494	TGGGTCCCTCCCATAACAC	ACT

dbSNP rs#	Extend Primer	Term Mix
rs1957728	AGAAGCATGTGCTTATAACAATAA	CGT
rs1957729	GAAGCATGTGCTTATAACAATAAA	ACT
rs1957730	ACATGAGAGACTCTGAAGACT	ACT
rs1957731	GGGTGAGCTTTGGGATCAC	ACT
rs1998468	GGGCATAATTAATCCATGTTAG	ACT
rs1957732	GGCCAAGTTTACCTGCAAAC	ACT
rs1957733	TCTAATGTTAAAGAGAGGAGTTTA	ACG
rs2376322	GCGCCAAGGAAAGGCCAC	ACT
rs2889345	TCATTCTCACCCTTGATATCCA	ACT
rs1815267	AAAGGGCTACTATCAGTTTTGT	CGT
rs1957734	CTGCCTTATAATTCTAAAAAGGT	ACT
rs1957735	CTAAACTAAGAAATGTTTCCAC	CGT
rs1957736	TAATACTAAGGAGAGGGCTCCT	ACT
rs1957737	AGCCAAGGGTGTGGATGAG	ACT
rs1957738	CCTTAAAGGCTGCCTACAAAATA	ACT
rs1957739	CTGAGTGCTTTAGCTGGATTA	ACG
rs1957740	TTAAGCATCACACTGAGTTTGAG	ACT
rs1957741	AGCTGAATTAAGCGCGACAGCTA	ACG
rs1957742	TCTACTTTGTACGTAGCTGTCGC	ACT
rs1957743	GAAAATATTACTAAAAAAGACCTC	ACG
rs1957744	TGTACCCTGTAATGCCTAAAGC	ACG
rs1957745	TTTTCAAAGGTTTAGGTTTGTTTT	ACT
rs1957746	ACAACATAGTTAGCAAATGCAG	ACG
rs1957747	GATAACAGTTCCAATTACAACAA	ACG
rs2146670	ATCTTTATGAGCTTTTCCTTTCTT	ACG
rs2146671	TACAACCCTTTCAGGACTTCA	CGT
rs1957748	TTGTAGTAGGGAGCCATGGT	ACT
rs2162307	CCTGGCCCTTTGTCCCTG	ACG
rs1962839	CCACATCTTTGACAAACCTGA	ACT
rs2376315	CCCCCTTCCTTTCCAGGC	ACT
rs1426410	CATCTTGCCCTAAAATCACTC	ACG
rs1895921	GTACATTTCTCAGGCAGCTC	ACG
rs1895922	ATTAGGCTTCTCCCACCATC	ACT
rs1035779	ACCCGGGAGGGTTGCAGT	ACT
rs1035780	GGCTGGAGTGCAGTGGCA	ACG
rs1035781	ACCTAGACTAAGAGAGTGATTGCA	ACT
rs1035782	CCTAAACAGTCAAGGCAAAGG	ACT
rs1426411	TTTATGGTCTTCTTAGGATATCA	ACG
rs1834602	AGGAAGGTGCCAGATCCT	ACG
rs1834603	AGGAAGGTGCCAGATCCTT	ACT
rs1834604	AGTTTTCTAGTAACTTCTCTAAAA	ACT
rs1834605	ACAGTCAGGAAGAATTCTGTCT	ACT
rs2162308	CACCTACAGAGTTTAAGTAAATTT	ACG
rs1365341	AAATCTCCTGGAGGGCTTCATAA	ACT

dbSNP rs#	Extend Primer	Term Mix
rs1820458	TGGAAATGGCAACTGAATCCT	ACT
rs1469310	ACCCACACAATGCCAATAGCAC	ACT
rs3057879	TGGAAAATAAAGCCTTTTGAGGTT	ACT
rs1469311	TGCCGTTAAAGAGGAAAAGCT	ACT
rs768326	CAGCTACTCTGTAAAGCTGAA	ACT
rs1863523	ATATTCTTGCTCATCTTTCTCTAT	ACT
rs1469312	TAGTCCAGCAAACGCCAGC	ACT
rs1469313	GTGAACAAATAATGCAAGTTCAG	ACT
rs1951773	CCCTTTGGGAGAGAAGGGC	ACT
rs2120655	AGCAATCCTCCCACCTTTGGC	CGT
rs2181495	GGTGACATTTGGGTGGGATACA	ACT

Genetic Analysis

[0235] Allelotyping results from the discovery cohort are shown for cases and controls in Table 17. The allele frequency for the A2 allele is noted in the fifth and sixth columns for osteoarthritis case pools and control pools, respectively, where “AF” is allele frequency. The allele frequency for the A1 allele can be easily calculated by subtracting the A2 allele frequency from 1 (A1 AF = 1-A2 AF). For example, the SNP rs1444311 has the following case and control allele frequencies: case A1 (A) = 0.74; case A2 (G) = 0.26; control A1 (A) = 0.75; and control A2 (G) = 0.25, where the nucleotide is provided in paranthesis. Some SNPs are labeled “untyped” because of failed assays.

TABLE 17

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs3849023	211	36870611	G/T			
rs1444311	7217	36877617	A/G	0.26	0.25	0.566
rs2044295	7895	36878295	A/C			
rs2166093	13308	36883708	C/T			
rs2376334	14279	36884679	G/T	0.15	0.16	0.734
rs1444320	17026	36887426	C/T			
rs2044294	18271	36888671	A/G	0.16	0.14	0.412
rs1899864	20417	36890817	C/T			
rs1562094	21843	36892243	A/G	0.22	0.23	0.586
rs1562098	22069	36892469	A/G			
rs1562097	22145	36892545	A/G	NA	0.97	NA
rs1562096	22519	36892919	A/G	0.20	0.21	0.773
rs1562095	22539	36892939	A/G	0.53	0.51	0.407
rs1444319	23236	36893636	A/C	0.74	0.79	0.023
rs1444318	23256	36893656	A/C	0.12	0.13	0.559
rs1025938	23402	36893802	C/T	0.18	0.19	0.633
rs1025937	23499	36893899	A/C			
rs1025936	23620	36894020	C/T	0.84	0.84	0.907
rs1020333	23871	36894271	A/T			
rs2120654	24136	36894536	C/G	0.15	0.16	0.718
rs2588547	25427	36895827	A/G	0.39	0.40	0.603
rs2044293	25866	36896266	G/T			

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs2760324	26541	36896941	A/G	0.59	0.61	0.395
rs2588546	26576	36896976	G/T	0.07	0.05	0.352
rs2588545	26689	36897089	A/G			
rs2760328	26720	36897120	A/C	0.25	0.26	0.791
rs2588544	27113	36897513	C/T			
rs2760331	27164	36897564	C/T	0.91	0.94	0.184
rs2588543	27186	36897586	A/G	0.59	0.59	0.828
rs2588542	28341	36898741	A/T			
rs2588541	29160	36899560	C/T	0.61	0.59	0.313
rs2588540	29844	36900244	A/G	0.62	0.62	0.999
rs2760336	30665	36901065	G/T			
rs2760337	30830	36901230	A/G	0.16	0.16	0.826
rs2028732	31061	36901461	A/C	0.60	0.58	0.432
rs2588538	31523	36901923	C/T	0.62	0.61	0.853
rs1992617	32326	36902726	C/T	0.61	0.59	0.282
rs1998469	32346	36902746	A/G			
rs1998470	32358	36902758	C/T	0.81	0.86	0.018
rs1975498	34909	36905309	C/T			
rs1562093	34975	36905375	A/G	0.89	0.87	0.529
rs1975497	35066	36905466	C/T	0.13	0.13	0.691
rs1562092	35096	36905496	G/T			
rs2248788	35375	36905775	C/T	0.29	0.31	0.368
rs1899862	36304	36906704	A/G	0.18	0.16	0.274
rs2588532	36712	36907112	A/T	0.30	0.32	0.443
rs1885878	36770	36907170	C/T	0.35	0.35	0.866
rs986648	37342	36907742	C/T	0.74	0.73	0.679
rs986647	37412	36907812	C/T	0.78	0.76	0.263
rs1010010	37884	36908284	A/G	0.25	0.26	0.649
rs1010009	38077	36908477	A/C	0.26	0.25	0.781
rs2760325	38300	36908700	C/T			
rs2588531	38301	36908701	C/T			
rs1838388	41189	36911589	C/T	0.75	0.74	0.650
rs1975495	44408	36914808	C/T			
rs2181491	44493	36914893	A/C	0.14	0.12	0.235
rs1975496	44571	36914971	A/G	0.26	0.26	0.944
rs2181492	44670	36915070	A/G	0.11	0.09	0.311
rs2224719	45219	36915619	A/G	0.78	0.78	0.866
rs2224720	45258	36915658	C/T	0.20	0.21	0.641
rs1951770	47261	36917661	A/G	0.22	0.18	0.029
rs2296040	48473	36918873	A/C	0.41	0.43	0.459
rs1957723	48771	36919171	A/G	0.42	0.38	0.113
rs1957725	55292	36925692	C/T	0.75	0.78	0.196
rs2889346	56479	36926879	A/G	0.54	0.55	0.677
rs1885879	56747	36927147	A/C	0.44	0.48	0.123
rs1957726	60620	36931020	G/T	0.14	0.14	0.741
rs1957727	60688	36931088	A/C	0.73	0.76	0.271
rs1885880	61058	36931458	A/C	0.43	0.43	0.935
rs1885881	61129	36931529	C/T	0.12	0.11	0.681
rs942108	61577	36931977	C/T	0.49	0.52	0.317
rs1951771	61961	36932361	A/G	0.93	NA	NA
rs2376323	63351	36933751	G/T			
rs2013358	63926	36934326	A/G	0.13	0.13	0.821
rs2181494	65798	36936198	A/G	0.42	0.43	0.512
rs1957728	66043	36936443	A/C			
rs1957729	66044	36936444	A/G	0.79	0.77	0.405
rs1957730	66246	36936646	C/T	0.15	0.16	0.719
rs1957731	66318	36936718	C/T	0.14	0.16	0.413

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs1998468	66547	36936947	G/T	0.13	0.12	0.468
rs1957732	71238	36941638	C/T	0.10	0.10	0.841
rs1957733	71283	36941683	A/G	0.63	0.61	0.632
rs2376322	71492	36941892	A/G	0.26	0.28	0.509
rs2889345	72274	36942674	A/G	0.20	0.18	0.234
rs1815267	73762	36944162	A/T	0.46	0.45	0.674
rs1957734	74209	36944609	G/T	0.55	0.64	0.003
rs1957735	75284	36945684	A/T	0.63	0.61	0.430
rs1957736	77347	36947747	A/C	0.05	0.05	0.903
rs1957737	77589	36947989	C/T	0.71	0.75	0.164
rs1957738	78096	36948496	A/G			
rs1957739	78606	36949006	A/G			
rs1957740	78862	36949262	G/T			
rs1957741	79135	36949535	A/G	0.76	0.80	0.077
rs1957742	79146	36949546	A/G	0.95	0.96	0.500
rs1957743	79456	36949856	C/T	0.21	0.16	0.039
rs1957744	79609	36950009	A/G	0.66	0.70	0.088
rs1957745	80086	36950486	A/G	0.88	0.90	0.354
rs1957746	80119	36950519	A/G	0.40	0.44	0.120
rs1957747	80766	36951166	C/T	0.72	0.76	0.093
rs2146670	81110	36951510	A/G	0.73	0.77	0.072
rs2146671	81269	36951669	A/T	0.17	0.15	0.250
rs1957748	81668	36952068	C/T	0.16	0.14	0.407
rs2162307	82433	36952833	C/T	0.73	0.76	0.170
rs1962839	82559	36952959	C/G			
rs2376315	83298	36953698	C/T	0.62	0.66	0.179
rs1426410	83821	36954221	A/G	0.75	0.77	0.307
rs1895921	84121	36954521	C/T	0.75	0.78	0.175
rs1895922	84147	36954547	C/T	0.15	0.12	0.095
rs1035779	84543	36954943	A/G	0.66	0.64	0.649
rs1035780	84554	36954954	A/G			
rs1035781	84691	36955091	A/G	0.73	0.77	0.100
rs1035782	84727	36955127	A/G			
rs1426411	85678	36956078	C/T	0.76	0.80	0.084
rs1834602	86699	36957099	C/T	0.20	0.16	0.072
rs1834603	86700	36957100	A/G	0.94	0.92	0.326
rs1834604	86792	36957192	A/G	0.70	0.73	0.287
rs1834605	86832	36957232	A/G	0.72	0.76	0.057
rs2162308	87045	36957445	A/G			
rs1365341	87140	36957540	A/G	0.18	0.15	0.086
rs1820458	87365	36957765	A/C	0.23	0.21	0.298
rs1469310	88342	36958742	C/T	0.20	0.18	0.265
rs3057879	88498	36958898	-/TCA	0.70	0.71	0.649
rs1469311	88589	36958989	A/G	0.70	0.74	0.065
rs768326	95502	36965902	A/G			
rs1863523	96968	36967368	C/T	0.21	0.18	0.247
rs1469312	97448	36967848	C/T	0.78	0.76	0.312
rs1469313	97568	36967968	C/T	0.81	0.80	0.617
rs1951773	98724	36969124	C/T			
rs2120655	Not mapped	Not mapped	T/G			
rs2181495	Not mapped	Not mapped	G/A	0.78	0.76	0.617

[0236] The *chrom 4* proximal SNPs were also allelotyped in the replication cohorts using the methods described herein and the primers provided in Tables 15 and 16. The replication allelotyping results for replication cohort #1 and replication cohort #2 are provided in Tables 18 and 19, respectively.

TABLE 18

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	A2 Case AF	A2 Control AF	p-Value
rs3849023	211	36870611	G/T			
rs1444311	7217	36877617	A/G	0.27	0.25	0.441
rs2044295	7895	36878295	A/C			
rs2166093	13308	36883708	C/T			
rs2376334	14279	36884679	G/T	0.14	0.14	0.970
rs1444320	17026	36887426	C/T			
rs2044294	18271	36888671	A/G	0.16	0.14	0.423
rs1899864	20417	36890817	C/T			
rs1562094	21843	36892243	A/G	0.22	0.21	0.725
rs1562098	22069	36892469	A/G			
rs1562097	22145	36892545	A/G			
rs1562096	22519	36892919	A/G	0.20	0.19	0.795
rs1562095	22539	36892939	A/G	0.55	0.53	0.453
rs1444319	23236	36893636	A/C	0.70	0.80	0.003
rs1444318	23256	36893656	A/C	0.12	0.13	0.645
rs1025938	23402	36893802	C/T	0.18	0.18	0.824
rs1025937	23499	36893899	A/C			
rs1025936	23620	36894020	C/T	0.85	0.83	0.622
rs1020333	23871	36894271	A/T			
rs2120654	24136	36894536	C/G	0.16	0.16	0.914
rs2588547	25427	36895827	A/G	0.40	0.40	0.980
rs2044293	25866	36896266	G/T			
rs2760324	26541	36896941	A/G	0.57	0.61	0.287
rs2588546	26576	36896976	G/T	0.08	0.05	0.265
rs2588545	26689	36897089	A/G			
rs2760328	26720	36897120	A/C	0.25	untyped	NA
rs2588544	27113	36897513	C/T			
rs2760331	27164	36897564	C/T	0.88	0.92	0.193
rs2588543	27186	36897586	A/G	0.57	0.58	0.869
rs2588542	28341	36898741	A/T			
rs2588541	29160	36899560	C/T	0.61	0.57	0.230
rs2588540	29844	36900244	A/G	0.64	0.64	0.926
rs2760336	30665	36901065	G/T			
rs2760337	30830	36901230	A/G	0.16	0.16	0.956
rs2028732	31061	36901461	A/C	0.60	0.57	0.330
rs2588538	31523	36901923	C/T	0.62	0.61	0.747
rs1992617	32326	36902726	C/T	0.62	0.59	0.341
rs1998469	32346	36902746	A/G			
rs1998470	32358	36902758	C/T	0.78	0.88	~0.0001
rs1975498	34909	36905309	C/T			
rs1562093	34975	36905375	A/G	0.89	0.90	0.905
rs1975497	35066	36905466	C/T	0.12	0.12	0.873
rs1562092	35096	36905496	G/T			
rs2248788	35375	36905775	C/T	0.28	0.31	0.308
rs1899862	36304	36906704	A/G	0.19	0.14	0.088
rs2588532	36712	36907112	A/T	0.30	0.33	0.347
rs1885878	36770	36907170	C/T	0.36	0.34	0.362
rs986648	37342	36907742	C/T	0.74	0.75	0.773
rs986647	37412	36907812	C/T	0.78	0.77	0.693

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	A2 Case AF	A2 Control AF	p-Value
rs1010010	37884	36908284	A/G	0.25	0.26	0.690
rs1010009	38077	36908477	A/C	0.27	0.26	0.870
rs2760325	38300	36908700	C/T			
rs2588531	38301	36908701	C/T			
rs1838388	41189	36911589	C/T	0.74	0.75	0.826
rs1975495	44408	36914808	C/T			
rs2181491	44493	36914893	A/C	0.16	0.10	0.057
rs1975496	44571	36914971	A/G	0.25	0.26	0.596
rs2181492	44670	36915070	A/G	0.11	0.08	0.167
rs2224719	45219	36915619	A/G	0.78	0.79	0.705
rs2224720	45258	36915658	C/T	0.19	0.21	0.478
rs1951770	47261	36917661	A/G	0.25	0.16	0.003
rs2296040	48473	36918873	A/C	0.39	0.43	0.241
rs1957723	48771	36919171	A/G	0.44	0.36	0.027
rs1957725	55292	36925692	C/T	0.73	0.80	0.024
rs2889346	56479	36926879	A/G	0.53	0.55	0.552
rs1885879	56747	36927147	A/C	0.43	0.50	0.057
rs1957726	60620	36931020	G/T	0.14	0.14	0.918
rs1957727	60688	36931088	A/C	0.71	0.78	0.038
rs1885880	61058	36931458	A/C	0.44	0.42	0.627
rs1885881	61129	36931529	C/T	0.12	0.12	0.833
rs942108	61577	36931977	C/T	0.42	0.49	0.096
rs1951771	61961	36932361	A/G	0.93	NA	NA
rs2376323	63351	36933751	G/T			
rs2013358	63926	36934326	A/G	0.13	0.12	0.795
rs2181494	65798	36936198	A/G	0.38	0.41	0.424
rs1957728	66043	36936443	A/C			
rs1957729	66044	36936444	A/G	0.78	0.77	0.672
rs1957730	66246	36936646	C/T	0.15	0.15	0.885
rs1957731	66318	36936718	C/T	0.15	0.16	0.719
rs1998468	66547	36936947	G/T	0.14	0.10	0.243
rs1957732	71238	36941638	C/T	0.10	0.09	0.817
rs1957733	71283	36941683	A/G	0.62	NA	0.628
rs2376322	71492	36941892	A/G	0.26	0.27	0.660
rs2889345	72274	36942674	A/G	0.22	0.16	0.020
rs1815267	73762	36944162	A/T	0.46	0.48	0.626
rs1957734	74209	36944609	G/T	0.44	0.61	~0.0001
rs1957735	75284	36945684	A/T	0.63	0.63	0.792
rs1957736	77347	36947747	A/C	0.03	0.03	0.987
rs1957737	77589	36947989	C/T	0.69	0.77	0.024
rs1957738	78096	36948496	A/G			
rs1957739	78606	36949006	A/G			
rs1957740	78862	36949262	G/T			
rs1957741	79135	36949535	A/G	0.75	0.83	0.008
rs1957742	79146	36949546	A/G	0.94	0.96	0.459
rs1957743	79456	36949856	C/T	0.24	0.14	0.009
rs1957744	79609	36950009	A/G	0.63	0.72	0.006
rs1957745	80086	36950486	A/G	0.86	0.90	0.229
rs1957746	80119	36950519	A/G	0.42	0.50	0.019
rs1957747	80766	36951166	C/T	0.71	0.79	0.009
rs2146670	81110	36951510	A/G	0.72	0.81	0.004
rs2146671	81269	36951669	A/T	0.17	0.13	0.106
rs1957748	81668	36952068	C/T	0.17	0.13	0.133
rs2162307	82433	36952833	C/T	0.72	0.78	0.020
rs1962839	82559	36952959	C/G			
rs2376315	83298	36953698	C/T	0.61	0.67	0.074
rs1426410	83821	36954221	A/G	0.73	0.79	0.058

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	A2 Case AF	A2 Control AF	p-Value
rs1895921	84121	36954521	C/T	0.72	0.80	0.013
rs1895922	84147	36954547	C/T	0.17	0.11	0.014
rs1035779	84543	36954943	A/G	0.66	0.64	0.613
rs1035780	84554	36954954	A/G			
rs1035781	84691	36955091	A/G	0.71	0.78	0.059
rs1035782	84727	36955127	A/G			
rs1426411	85678	36956078	C/T	0.75	0.82	0.008
rs1834602	86699	36957099	C/T	0.22	0.15	0.020
rs1834603	86700	36957100	A/G	0.94	0.92	0.483
rs1834604	86792	36957192	A/G	0.69	0.75	0.056
rs1834605	86832	36957232	A/G	0.71	0.79	0.007
rs2162308	87045	36957445	A/G			
rs1365341	87140	36957540	A/G	0.19	0.13	0.017
rs1820458	87365	36957765	A/C	0.24	0.19	0.141
rs1469310	88342	36958742	C/T	0.22	0.17	0.061
rs3057879	88498	36958898	-/TCA	0.67	NA	NA
rs1469311	88589	36958989	A/G	0.67	0.76	0.006
rs768326	95502	36965902	A/G			
rs1863523	96968	36967368	C/T	0.22	0.17	0.103
rs1469312	97448	36967848	C/T	0.80	0.77	0.236
rs1469313	97568	36967968	C/T	0.83	0.80	0.422
rs1951773	98724	36969124	C/T			
rs2120655	Not mapped	Not mapped	T/G			
rs2181495	Not mapped	Not mapped	G/A	0.78	0.76	0.617

TABLE 19

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	A2 Case AF	A2 Control AF	p-Value
rs3849023	211	36870611	G/T			
rs1444311	7217	36877617	A/G	0.25	0.26	0.876
rs2044295	7895	36878295	A/C			
rs2166093	13308	36883708	C/T			
rs2376334	14279	36884679	G/T	0.16	0.18	0.532
rs1444320	17026	36887426	C/T			
rs2044294	18271	36888671	A/G	NA	0.15	NA
rs1899864	20417	36890817	C/T			
rs1562094	21843	36892243	A/G	NA	0.28	NA
rs1562098	22069	36892469	A/G			
rs1562097	22145	36892545	A/G			
rs1562096	22519	36892919	A/G	0.20	0.23	0.364
rs1562095	22539	36892939	A/G	0.50	0.48	0.588
rs1444319	23236	36893636	A/C	0.79	0.79	0.923
rs1444318	23256	36893656	A/C	0.12	0.13	0.711
rs1025938	23402	36893802	C/T	0.18	0.22	0.247
rs1025937	23499	36893899	A/C			
rs1025936	23620	36894020	C/T	0.84	0.86	0.403
rs1020333	23871	36894271	A/T			
rs2120654	24136	36894536	C/G	0.14	0.16	0.682
rs2588547	25427	36895827	A/G	0.37	NA	
rs2044293	25866	36896266	G/T			
rs2760324	26541	36896941	A/G	0.60	0.60	0.965
rs2588546	26576	36896976	G/T	0.05	0.06	0.797
rs2588545	26689	36897089	A/G			
rs2760328	26720	36897120	A/C	0.25	0.26	0.816

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	A2 Case AF	A2 Control AF	p-Value
rs2588544	27113	36897513	C/T			
rs2760331	27164	36897564	C/T	0.95	0.96	0.597
rs2588543	27186	36897586	A/G	0.60	0.61	0.801
rs2588542	28341	36898741	A/T			
rs2588541	29160	36899560	C/T	0.62	0.61	0.972
rs2588540	29844	36900244	A/G	0.60	0.59	0.810
rs2760336	30665	36901065	G/T			
rs2760337	30830	36901230	A/G	0.16	0.17	0.659
rs2028732	31061	36901461	A/C	0.60	0.60	0.976
rs2588538	31523	36901923	C/T	0.61	0.61	0.912
rs1992617	32326	36902726	C/T	0.61	0.59	0.583
rs1998469	32346	36902746	A/G			
rs1998470	32358	36902758	C/T	0.84	0.81	0.338
rs1975498	34909	36905309	C/T			
rs1562093	34975	36905375	A/G	0.88	0.84	0.199
rs1975497	35066	36905466	C/T	0.13	0.15	0.613
rs1562092	35096	36905496	G/T			
rs2248788	35375	36905775	C/T	0.30	0.31	0.884
rs1899862	36304	36906704	A/G	0.17	0.18	0.563
rs2588532	36712	36907112	A/T	0.29	0.29	0.952
rs1885878	36770	36907170	C/T	0.33	0.36	0.405
rs986648	37342	36907742	C/T	0.75	0.72	0.283
rs986647	37412	36907812	C/T	0.79	0.74	0.186
rs1010010	37884	36908284	A/G	0.25	0.26	0.843
rs1010009	38077	36908477	A/C	0.25	0.24	0.764
rs2760325	38300	36908700	C/T			
rs2588531	38301	36908701	C/T			
rs1838388	41189	36911589	C/T	0.76	0.72	0.284
rs1975495	44408	36914808	C/T			
rs2181491	44493	36914893	A/C	0.12	0.15	0.464
rs1975496	44571	36914971	A/G	0.27	0.25	0.577
rs2181492	44670	36915070	A/G	0.10	0.11	0.844
rs2224719	45219	36915619	A/G	0.78	0.75	0.426
rs2224720	45258	36915658	C/T	0.21	0.20	0.790
rs1951770	47261	36917661	A/G	0.19	0.20	0.796
rs2296040	48473	36918873	A/C	0.43	0.42	0.804
rs1957723	48771	36919171	A/G	0.41	0.42	0.653
rs1957725	55292	36925692	C/T	0.77	0.75	0.439
rs2889346	56479	36926879	A/G	0.56	0.56	0.948
rs1885879	56747	36927147	A/C	0.46	0.45	0.959
rs1957726	60620	36931020	G/T	0.14	0.15	0.673
rs1957727	60688	36931088	A/C	0.76	0.73	0.255
rs1885880	61058	36931458	A/C	0.43	0.45	0.614
rs1885881	61129	36931529	C/T	0.13	0.10	0.346
rs942108	61577	36931977	C/T	0.58	0.56	0.730
rs1951771	61961	36932361	A/G			
rs2376323	63351	36933751	G/T			
rs2013358	63926	36934326	A/G	0.13	0.15	0.469
rs2181494	65798	36936198	A/G	0.46	0.47	0.820
rs1957728	66043	36936443	A/C			
rs1957729	66044	36936444	A/G	0.80	0.78	0.440
rs1957730	66246	36936646	C/T	0.15	0.17	0.668
rs1957731	66318	36936718	C/T	0.14	0.16	0.387
rs1998468	66547	36936947	G/T	0.12	0.13	0.615
rs1957732	71238	36941638	C/T	0.09	0.11	0.469
rs1957733	71283	36941683	A/G	0.60	0.02	
rs2376322	71492	36941892	A/G	0.27	0.29	0.582

dbSNP rs#	Position in SEQ ID NO: 2	Chromosome Position	A1/A2 Allele	A2 Case AF	A2 Control AF	p-Value
rs2889345	72274	36942674	A/G	0.17	0.21	0.308
rs1815267	73762	36944162	A/T	0.46	0.41	0.151
rs1957734	74209	36944609	G/T	0.68	0.69	0.766
rs1957735	75284	36945684	A/T	0.62	0.58	0.311
rs1957736	77347	36947747	A/C	0.07	0.08	0.688
rs1957737	77589	36947989	C/T	0.75	0.71	0.305
rs1957738	78096	36948496	A/G			
rs1957739	78606	36949006	A/G			
rs1957740	78862	36949262	G/T			
rs1957741	79135	36949535	A/G	0.78	0.76	0.446
rs1957742	79146	36949546	A/G	0.96	0.96	0.938
rs1957743	79456	36949856	C/T	0.17	0.19	0.667
rs1957744	79609	36950009	A/G	0.69	0.66	0.423
rs1957745	80086	36950486	A/G	0.90	0.89	0.738
rs1957746	80119	36950519	A/G	0.37	0.35	0.708
rs1957747	80766	36951166	C/T	0.72	0.70	0.639
rs2146670	81110	36951510	A/G	0.75	0.72	0.306
rs2146671	81269	36951669	A/T	0.16	0.17	0.806
rs1957748	81668	36952068	C/T	0.14	0.17	0.453
rs2162307	82433	36952833	C/T	0.76	0.73	0.465
rs1962839	82559	36952959	C/G			
rs2376315	83298	36953698	C/T	0.64	0.63	0.767
rs1426410	83821	36954221	A/G	0.77	0.74	0.465
rs1895921	84121	36954521	C/T	0.78	0.75	0.320
rs1895922	84147	36954547	C/T	0.12	0.13	0.586
rs1035779	84543	36954943	A/G	NA	0.65	NA
rs1035780	84554	36954954	A/G			
rs1035781	84691	36955091	A/G	0.75	0.76	0.830
rs1035782	84727	36955127	A/G			
rs1426411	85678	36956078	C/T	0.78	0.76	0.488
rs1834602	86699	36957099	C/T	0.18	0.18	0.945
rs1834603	86700	36957100	A/G	0.95	NA	
rs1834604	86792	36957192	A/G	0.72	0.69	0.427
rs1834605	86832	36957232	A/G	0.73	0.72	0.647
rs2162308	87045	36957445	A/G			
rs1365341	87140	36957540	A/G	0.16	0.17	0.667
rs1820458	87365	36957765	A/C	0.23	0.24	0.670
rs1469310	88342	36958742	C/T	0.19	0.21	0.592
rs3057879	88498	36958898	-/TCA	0.74	0.71	0.478
rs1469311	88589	36958989	A/G	0.74	0.72	0.582
rs768326	95502	36965902	A/G			
rs1863523	96968	36967368	C/T	0.20	0.21	0.687
rs1469312	97448	36967848	C/T	0.76	0.75	0.807
rs1469313	97568	36967968	C/T	0.78	0.79	0.824
rs1951773	98724	36969124	C/T			
rs2120655	Not mapped	Not mapped	T/G			
rs2181495	Not mapped	Not mapped	G/A			

[0237] Allelotyping results were considered particularly significant with a calculated p-value of less than or equal to 0.05 for allelotype results. These values are indicated in bold. The allelotyping p-values were plotted in Figure 1B for the discovery cohort. The position of each SNP on the chromosome is presented on the x-axis. The y-axis gives the negative logarithm (base 10) of the p-value comparing the

estimated allele in the case group to that of the control group. The minor allele frequency of the control group for each SNP designated by an X or other symbol on the graphs in Figure 1B can be determined by consulting Table 17. For example, the left-most X on the left graph is at position 36870611. By proceeding down the Table from top to bottom and across the graphs from left to right the allele frequency associated with each symbol shown can be determined.

[0238] To aid the interpretation, multiple lines have been added to the graph. The broken horizontal lines are drawn at two common significance levels, 0.05 and 0.01. The vertical broken lines are drawn every 20kb to assist in the interpretation of distances between SNPs. Two other lines are drawn to expose linear trends in the association of SNPs to the disease. The light gray line (or generally bottom-most curve) is a nonlinear smoother through the data points on the graph using a local polynomial regression method (W.S. Cleveland, E. Grosse and W.M. Shyu (1992) Local regression models. Chapter 8 of Statistical Models in S eds J.M. Chambers and T.J. Hastie, Wadsworth & Brooks/Cole). The black line provides a local test for excess statistical significance to identify regions of association. This was created by use of a 10kb sliding window with 1kb step sizes. Within each window, a chi-square goodness of fit test was applied to compare the proportion of SNPs that were significant at a test wise level of 0.01, to the proportion that would be expected by chance alone (0.05 for the methods used here). Resulting p-values that were less than 10^{-8} were truncated at that value.

[0239] Finally, the exons and introns of the genes in the covered region are plotted below each graph at the appropriate chromosomal positions. The gene boundary is indicated by the broken horizontal line. The exon positions are shown as thick, unbroken bars. An arrow is placed at the 3' end of each gene to show the direction of transcription.

Example 6 SNW1 Region Proximal SNPs

[0240] SNP rs1477261 is associated with osteoarthritis and is described in Table A. It lies within an intron of the SKI-interacting protein gene (SNW1). This gene, a member of the SNW gene family, encodes a coactivator that enhances transcription from some Pol II promoters. This coactivator can bind to the ligand-binding domain of the vitamin D receptor and to retinoid receptors to enhance vitamin D-, retinoic acid-, estrogen-, and glucocorticoid-mediated gene expression. It also can interact with poly(A)-binding protein 2 to directly control the expression of muscle-specific genes at the transcriptional level. One hundred sixty-three additional allelic variants proximal to rs1477261 were identified and subsequently allelotyped in osteoarthritis case and control sample sets as described in Examples 1 and 2. The polymorphic variants are set forth in Table 20. The chromosome position provided in column four of Table 20 is based on Genome "Build 34" of NCBI's GenBank.

TABLE 20

dbSNP rs#	Chromosome	Position in SEQ ID NO: 3	Chromosome Position	Allele Variants
rs7143926	14	218	76161268	a/t
rs1549071	14	1440	76162490	c/t
rs8012858	14	1442	76162492	c/t
rs7155611	14	2611	76163661	c/t
rs176941	14	4317	76165367	a/c
rs176942	14	4724	76165774	a/g
rs176943	14	4788	76165838	g/t
rs176944	14	5202	76166252	g/t
rs4365221	14	5780	76166830	c/t
rs3168952	14	5974	76167024	c/t
rs176945	14	6644	76167694	c/g
rs176946	14	7430	76168480	a/g
rs176947	14	7938	76168988	c/t
rs176948	14	8095	76169145	c/t
rs176949	14	8183	76169233	a/c
rs176950	14	8312	76169362	c/t
rs176951	14	8352	76169402	a/c
rs7156905	14	9348	76170398	c/t
rs3217197	14	9378	76170428	-/tctc
rs2270443	14	9617	76170667	a/g
rs176952	14	9727	76170777	c/t
rs176953	14	9834	76170884	c/t
rs176954	14	9899	76170949	g/t
rs176955	14	10211	76171261	c/t
rs3214416	14	10377	76171427	-/t
rs176956	14	10695	76171745	c/t
rs2544566	14	10729	76171779	c/g
rs2544567	14	10730	76171780	c/t
rs176957	14	11433	76172483	a/g
rs176958	14	11951	76173001	c/g
rs176959	14	12697	76173747	c/t
rs1802227	14	12982	76174032	a/c
rs176961	14	14419	76175469	c/t
rs176962	14	14501	76175551	c/t
rs7401285	14	14983	76176033	a/c
rs176963	14	15280	76176330	c/t
rs176964	14	15475	76176525	a/g
rs4903631	14	15888	76176938	a/g
rs4903632	14	15976	76177026	a/t
rs176965	14	16307	76177357	a/c
rs4903633	14	16442	76177492	a/c
rs176966	14	17255	76178305	c/t
rs176968	14	18948	76179998	g/t
rs176969	14	19435	76180485	a/t
rs176970	14	19753	76180803	c/t
rs7149198	14	20021	76181071	c/t

dbSNP rs#	Chromosome	Position in SEQ ID NO: 3	Chromosome Position	Allele Variants
rs7147918	14	20022	76181072	a/c
rs7148685	14	20503	76181553	a/g
rs1184232	14	20590	76181640	g/t
rs1184233	14	21804	76182854	g/t
rs1184234	14	21919	76182969	c/t
rs7401998	14	21990	76183040	a/t
rs176974	14	22412	76183462	a/g
rs6574390	14	22536	76183586	c/t
rs176975	14	23432	76184482	a/g
rs176976	14	23468	76184518	g/t
rs176977	14	23772	76184822	c/t
rs8013727	14	24325	76185375	c/t
rs176978	14	24773	76185823	c/t
rs2111829	14	26274	76187324	c/t
rs176980	14	27440	76188490	c/g
rs5809848	14	28561	76189611	-/acag
rs5809849	14	30071	76191121	-/a
rs4383070	14	31764	76192814	a/t
rs7493652	14	33008	76194058	c/t
rs2112133	14	35310	76196360	a/t
rs1963833	14	35460	76196510	a/c
rs6574391	14	37112	76198162	a/g
rs7155062	14	37285	76198335	a/g
rs4899674	14	37747	76198797	c/t
rs8022516	14	38057	76199107	c/t
rs7140838	14	38859	76199909	a/c
rs7141127	14	38860	76199910	a/g
rs6574392	14	39525	76200575	a/g
rs8003691	14	40216	76201266	a/g
rs8003979	14	40281	76201331	c/t
rs8010541	14	41453	76202503	c/g
rs8016416	14	42091	76203141	a/t
rs8016175	14	42513	76203563	a/g
rs7154571	14	42935	76203985	c/t
rs7158826	14	42985	76204035	a/g
rs7159310	14	43003	76204053	a/g
rs7401900	14	43281	76204331	a/g
rs7160355	14	43716	76204766	c/t
rs2032781	14	43866	76204916	a/g
rs6574394	14	44234	76205284	g/t
rs8007598	14	44596	76205646	a/g
rs2267767	14	44871	76205921	c/t
rs6574395	14	45005	76206055	a/g
rs7150066	14	45282	76206332	a/c
rs7492334	14	47178	76208228	a/c
rs4359361	14	47816	76208866	g/t
rs4605089	14	47887	76208937	a/g

dbSNP rs#	Chromosome	Position in SEQ ID NO: 3	Chromosome Position	Allele Variants
rs7146446	14	48134	76209184	c/t
rs4346144	14	48135	76209185	a/g
rs7148078	14	48276	76209326	g/t
rs7148286	14	48400	76209450	c/t
rs3783980	14	48798	76209848	a/g
rs1549119	14	48803	76209853	a/t
rs1984925	14	49146	76210196	c/t
rs1477261	14	49969	76211019	a/t
rs8016447	14	51059	76212109	a/g
rs7494044	14	51064	76212114	c/t
rs2023288	14	53285	76214335	a/t
rs7151685	14	54560	76215610	c/t
rs2112135	14	54748	76215798	a/g
rs2161088	14	54785	76215835	c/g
rs4903638	14	55102	76216152	c/g
rs1477262	14	55644	76216694	a/g
rs1477263	14	55705	76216755	g/t
rs1477264	14	55841	76216891	a/g
rs2277917	14	56623	76217673	c/g
rs2277918	14	56825	76217875	a/c
rs2277919	14	56827	76217877	a/g
rs1978416	14	56892	76217942	c/t
rs3759728	14	59150	76220200	a/t
rs6574399	14	59958	76221008	a/t
rs7155336	14	60231	76221281	c/t
rs7156186	14	60524	76221574	a/g
rs7142390	14	61871	76222921	c/t
rs7145875	14	62226	76223276	c/t
rs8014635	14	63230	76224280	g/t
rs8015938	14	63468	76224518	g/t
rs8015313	14	63787	76224837	c/t
rs8006315	14	65732	76226782	a/c
rs6574400	14	65989	76227039	a/g
rs7140816	14	68832	76229882	g/t
rs4566078	14	69904	76230954	c/t
rs7141050	14	70365	76231415	a/g
rs3049356	14	70886	76231936	-/tatc
rs4903639	14	73088	76234138	a/t
rs4903641	14	73103	76234153	c/t
rs2364838	14	75934	76236984	c/t
rs2364839	14	75966	76237016	c/t
rs4632066	14	76273	76237323	c/t
rs2112136	14	77943	76238993	c/t
rs4641655	14	78466	76239516	c/t
rs4635269	14	78861	76239911	c/t
rs4570764	14	78872	76239922	a/g
rs759808	14	79836	76240886	g/t

dbSNP rs#	Chromosome	Position in SEQ ID NO: 3	Chromosome Position	Allele Variants
rs7150531	14	80908	76241958	c/t
rs7154968	14	81509	76242559	c/g
rs7146657	14	83576	76244626	c/t
rs7145859	14	83662	76244712	c/g
rs4903643	14	83782	76244832	c/t
rs717682	14	84282	76245332	g/t
rs717683	14	84444	76245494	a/g
rs1477259	14	85129	76246179	c/g
rs8019064	14	85151	76246201	a/g
rs8018971	14	85296	76246346	a/c
rs1477260	14	85809	76246859	c/g
rs5809851	14	86387	76247437	-/t
rs1985149	14	86494	76247544	a/g
rs1008988	14	89786	76250836	a/g
rs1008989	14	89894	76250944	a/t
rs8018222	14	90122	76251172	g/t
rs1006040	14	92067	76253117	a/g
rs1006039	14	92187	76253237	c/t
rs1006038	14	92312	76253362	a/g
rs8009784	14	92824	76253874	g/t
rs4903644	14	93733	76254783	c/t
rs7149496	14	96553	76257603	c/g
rs6574402	14	96941	76257991	a/c

Assay for Verifying and Allelotyping SNPs

[0241] The methods used to verify and allelotype the 101 proximal SNPs of Table 20 are the same methods described in Examples 1 and 2 herein. The primers and probes used in these assays are provided in Table 21 and Table 22, respectively.

TABLE 21

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs7143926	ACGTTGGATGGAGTCACCCAAAATTAAGGC	ACGTTGGATGGAAAGCCAAAATTAGCCTGC
rs1549071	ACGTTGGATGGTGAGACGCTGTCTCAGTAA	ACGTTGGATGCTCCACACTTGGAGAAGTTG
rs8012858	ACGTTGGATGGTGAGACGCTGTCTCAGTAA	ACGTTGGATGCTCCACACTTGGAGAAGTTG
rs7155611	ACGTTGGATGATGGAATACAGGCACCGTTC	ACGTTGGATGCCCTTCTTAATCTCCATGG
rs176941	ACGTTGGATGTTAGTATGGGAAAAGGGCTC	ACGTTGGATGCAACAATCCTATGAGTTGGG
rs176942	ACGTTGGATGAGTGGCTCAGATGTGAGTAG	ACGTTGGATGTGGTCTTACCAACCACATG
rs176943	ACGTTGGATGACCAAGCCAGTAAAGTCTC	ACGTTGGATGGCATCCGCAAGATGCTAATG
rs176944	ACGTTGGATGGGCCTCAATATTGGCTAAATG	ACGTTGGATGCTTAACCATTAGAGCCCTTC
rs4365221	ACGTTGGATGAAATAAGGCAGGAAGGGTAG	ACGTTGGATGTCCCAACTTACTGGTCTTTC
rs3168952	ACGTTGGATGATGTACCAGACTTGGTGGTG	ACGTTGGATGTTTGCTGAGGATGGAGACTG
rs176945	ACGTTGGATGCCTACTATACACTCACAAA	ACGTTGGATGTTTTTAAACACTTTAAGC

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs176946	ACGTTGGATGGCTTTATCATAGGTATTTGTG	ACGTTGGATGGAGAGATGTGTTGTTTTGAG
rs176947	ACGTTGGATGTGAGTAGCTGGGACTACAGG	ACGTTGGATGGGCCAACATAGCGAACTCC
rs176948	ACGTTGGATGCAGAGCCAAAGGTCAACAAG	ACGTTGGATGTACAGGTGTGAGCCTTCATG
rs176949	ACGTTGGATGTAGGAACTCCCTGCAGTTCC	ACGTTGGATGCCTTGCTGGCTTTAAAGAAG
rs176950	ACGTTGGATGAATCACAGGAGTGACATCCC	ACGTTGGATGTGGAGGAGAAACCTGACTTG
rs176951	ACGTTGGATGCCCTATATAATCTCCTCCCC	ACGTTGGATGCAGGAGTGACATCCCATTAC
rs7156905	ACGTTGGATGTGAGAGAGAGAACTGTCTC	ACGTTGGATGAAAGGCGGCTTTGATGTTGG
rs3217197	ACGTTGGATGTTGATTGTGCCACTGCACTC	ACGTTGGATGACTCTAGTTGAAAATCCTGG
rs2270443	ACGTTGGATGATAACTCAGTCCAGGTGTGG	ACGTTGGATGCACTCAAGCAGTCTACTCAC
rs176952	ACGTTGGATGGATCTCAGCTCACTGCAATC	ACGTTGGATGTATCTGGGTGACTGAGGAAG
rs176953	ACGTTGGATGTTGAGGTCAGGAGTTTGGGA	ACGTTGGATGGCCACCACACCCAGCTAATT
rs176954	ACGTTGGATGAAAACATAGGCCAGGTGCAG	ACGTTGGATGAAACTCCTGACCTCAAGCCA
rs176955	ACGTTGGATGCTAGAGTGCTTGGATGTACC	ACGTTGGATGGTCATCTACAGGGACTAGAC
rs3214416	ACGTTGGATGACGACTATCATCACGTGTTT	ACGTTGGATGACCAGAAGTCTGTAAGTAGG
rs176956	ACGTTGGATGTACAGGCATAAGCCACCATG	ACGTTGGATGAGGAAGGGTGTAAGCAAGG
rs2544566	ACGTTGGATGCAAGCAATCTTCCCATCTGG	ACGTTGGATGTGATCCGATTTTTGGCTGGG
rs2544567	ACGTTGGATGCAAGCAATCTTCCCATCTGG	ACGTTGGATGTGATCCGATTTTTGGCTGGG
rs176957	ACGTTGGATGTTTACCCTGTTAGCCAGGA	ACGTTGGATGTAATCCAGCACTTTGGGAG
rs176958	ACGTTGGATGAAAACCTGGGCACTCTACCAC	ACGTTGGATGAAAATCGCGCCATTGCACTC
rs176959	ACGTTGGATGCAGGCAGTTTTTATTTGTCCC	ACGTTGGATGGGTTAGGGAGTCATAATACC
rs1802227	ACGTTGGATGAACAAATAGTTGCACCAAG	ACGTTGGATGTTTTAATTTGGAGTGGGCA
rs176961	ACGTTGGATGAACCCAGTTTAAGACCGGCC	ACGTTGGATGTACAGGTGTGTGCCACCATG
rs176962	ACGTTGGATGATATTTCTGGCTGGGCACTG	ACGTTGGATGACTGGGTTCAAGCAATCTGC
rs7401285	ACGTTGGATGACAGAGTGGGACTCCATATC	ACGTTGGATGGATTCAAACCTGGGTGTCTTG
rs176963	ACGTTGGATGTAAGCCTGGGAAAACACACG	ACGTTGGATGCCCACTCTACTTTCCAGTAG
rs176964	ACGTTGGATGAGAGTCAGTGTCTACAAAA	ACGTTGGATGTAATCCCGTTTTACAGCTTC
rs4903631	ACGTTGGATGGTAAATGCCAGCATGATGAC	ACGTTGGATGTCTCAGCCCACTATAAGAAG
rs4903632	ACGTTGGATGTGTGAATACCTATCCTCAGG	ACGTTGGATGGTCATCATGCTGGCATTAC
rs176965	ACGTTGGATGAATGCTTTATAAGGGCTGCC	ACGTTGGATGTCTCAGAAACAAAGGATGTG
rs4903633	ACGTTGGATGCAACCCCCAAACCATCATAT	ACGTTGGATGCTAACAGATTCTGTTGACATGG
rs176966	ACGTTGGATGCTCTCGAGTAGCTGGGACTA	ACGTTGGATGTGGCCAACATGGTGAAACCC
rs176968	ACGTTGGATGGCGAAACTCCGTCTCAAAAC	ACGTTGGATGTAGTGATCTTCCACCTAGG
rs176969	ACGTTGGATGCTGTCTGTCCGATTTACTGC	ACGTTGGATGTCTAGAATCAAGCATGCGGC
rs176970	ACGTTGGATGCTAATGTTTCTTAGTACAGTGG	ACGTTGGATGCTTCTCTTCTAGCTATTTTGC
rs7149198	ACGTTGGATGCAATGGGATATTACTCAGCC	ACGTTGGATGTTTCTGTGCCGGGCTTATTC
rs7147918	ACGTTGGATGCAATGGGATATTACTCAGCC	ACGTTGGATGTTTCTGTGCCGGGCTTATTC
rs7148685	ACGTTGGATGTGTCTTCTTTGAGACCGTC	ACGTTGGATGCTCAATCGCAAAGAAACGAG
rs1184232	ACGTTGGATGAAGAGGGCCACCTACAGAATG	ACGTTGGATGCTCGTTTCTTTGCGATTGAG
rs1184233	ACGTTGGATGAAGTGTGGGATTACAGGTG	ACGTTGGATGAGTGAAAGATCGCCACAAAG
rs1184234	ACGTTGGATGGCTATGTGCAGTGAATCATG	ACGTTGGATGTCTCAGACCTCAGGTGATCT
rs7401998	ACGTTGGATGTGAGTAGCTAGGACAACAGG	ACGTTGGATGAACGTGGTGAAACCCCATCT
rs176974	ACGTTGGATGTTACAGCGAGCTGAGATCAT	ACGTTGGATGAGGATCATACTGTCTCTGAC
rs6574390	ACGTTGGATGTGATGAAACCCCGTCTGTAC	ACGTTGGATGTCCTGAGTAGCTGGGATTAC
rs176975	ACGTTGGATGTGTAGAATCTAGGTGGTAGG	ACGTTGGATGCCAGCCTTTCCTGACATTTT
rs176976	ACGTTGGATGGGTAGGAGATACAGGTGTTT	ACGTTGGATGCCAGCCTTTCCTGACATTT
rs176977	ACGTTGGATGTTGCATCATTACACTTCAGC	ACGTTGGATGGGGAAACATTATGCATAATTCC
rs8013727	ACGTTGGATGTGCCTGGTTGTATACCTAAC	ACGTTGGATGCTTGAGAACGATTCTGTTGTC

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs176978	ACGTTGGATGGGGACCATGTTTTGTTACC	ACGTTGGATGAATACTGTGGAATGGGCATG
rs2111829	ACGTTGGATGCATGTGGAAAAAGGTATGAC	ACGTTGGATGCCTACTTTATATGCAGTAGG
rs176980	ACGTTGGATGATGGCCAATGCTATGAACGC	ACGTTGGATGAAGGGCAGTTGCAGGAAAAAG
rs5809848	ACGTTGGATGTCTATTTTTCCAGAGCTTGGG	ACGTTGGATGCCATTTCACTGATGCTTTGG
rs5809849	ACGTTGGATGGTGAATACCGTGTCAGTTCC	ACGTTGGATGTGCAGTGAGCTGAGATCATG
rs4383070	ACGTTGGATGAGCGATTCTCTTGTCTCAGC	ACGTTGGATGAACTTAGCTGGGCATTGTGG
rs7493652	ACGTTGGATGGGTCAATATACCACAAGTAAC	ACGTTGGATGCTGGCCCTATGCTATTTTCA
rs2112133	ACGTTGGATGGCCACCACAAGTGGCTAATT	ACGTTGGATGTGTGGTCAGGAGATCGAAAC
rs1963833	ACGTTGGATGTAAGCCAAGATTGCGTCACT	ACGTTGGATGAGCATTAAAGGTAGAATGCC
rs6574391	ACGTTGGATGTAACCGTTGCTATGGAGAAG	ACGTTGGATGACCTATACAACCCTAAGCTG
rs7155062	ACGTTGGATGGCTCCTTATTTGGGCATTCC	ACGTTGGATGCACTCAGCCTTGTGAGATAC
rs4899674	ACGTTGGATGAATGTGCTGAGGAACTGAG	ACGTTGGATGGCTTCTGATACTTTCAAGAG
rs8022516	ACGTTGGATGGTTGAAGGCATTCTTTTGGG	ACGTTGGATGCTAGCCTGGGCAATATAATG
rs7140838	ACGTTGGATGCCTCGTTTCTGAAGAATACC	ACGTTGGATGGAGACTGAACAGGTTATTGG
rs7141127	ACGTTGGATGCCTCGTTTCTGAAGAATACC	ACGTTGGATGGAGACTGAACAGGTTATTGG
rs6574392	ACGTTGGATGAGAAAATAGCATAGGCTGGG	ACGTTGGATGAAATGATCCATCCTCCTCAG
rs8003691	ACGTTGGATGACTGAAGTCAAGTGAAGGCC	ACGTTGGATGTTAGGCCCTTATACATGGAG
rs8003979	ACGTTGGATGCACAAAACCACTTCTGAAGC	ACGTTGGATGGGGCCTAATTTTCTTTTGC
rs8010541	ACGTTGGATGCACTTTTCTTGGCTAGCTTC	ACGTTGGATGCAGAATGGCTAAAAGTGAAC
rs8016416	ACGTTGGATGTGCCATAACTTCTTTGAC	ACGTTGGATGGCCACGGAATCCTATATAGA
rs8016175	ACGTTGGATGTTGAGCACTGAGTGAGTGAG	ACGTTGGATGTCCTAACCGTGAGTGATCTG
rs7154571	ACGTTGGATGATGTGAGGAGCACCTCTGCC	ACGTTGGATGCTCTTCCCTTCTCAGACGG
rs7158826	ACGTTGGATGCACCTCCCTCCTGGACGGG	ACGTTGGATGGCCACCCCGTCTGAGAAGG
rs7159310	ACGTTGGATGACCCCGTCTGAGAAGGGAAG	ACGTTGGATGCACCTCCCTCCTGGACGGG
rs7401900	ACGTTGGATGCCCAACAGCTCATTGAGAAC	ACGTTGGATGTCTTTTCCCACATTTCCCC
rs7160355	ACGTTGGATGTCACCTGTTTATCTGCTGAC	ACGTTGGATGTTATTGATCATTCTTGGGTG
rs2032781	ACGTTGGATGTATATCACTGTAGTAACAGC	ACGTTGGATGACCATAAGTATATATCACAAG
rs6574394	ACGTTGGATGACCACACCCAGCCTATTTGT	ACGTTGGATGTTATGCTGAAAGCCTGGGAG
rs8007598	ACGTTGGATGCTGGCAAAAGTCTCTTAACAC	ACGTTGGATGTTGGTTAAAGTCACAGAATG
rs2267767	ACGTTGGATGGTTTCACCATGTTAGCCAGG	ACGTTGGATGTAATCCCAGCACTTTGGGAG
rs6574395	ACGTTGGATGAACCTTGAACCTTGGGGCTC	ACGTTGGATGAAAAAATTCACCGGGCATGG
rs7150066	ACGTTGGATGAAGCAATCCTCCTGCTTCTG	ACGTTGGATGAGATCAGGTGTAGATCCAGG
rs7492334	ACGTTGGATGGCCTTTGCATTGGCTATTTG	ACGTTGGATGTAGAAAGCAGTCATGGGAAG
rs4359361	ACGTTGGATGGTAGTATTTGCTTAGTACAC	ACGTTGGATGTTCTAAGCCTGAATGTTTCC
rs4605089	ACGTTGGATGAATACCTATGAGATCTCAGG	ACGTTGGATGCCTTGTAACCTTTAACATC
rs7146446	ACGTTGGATGATTCACTTTTACAAGACCTC	ACGTTGGATGGCATATTGTACTTAGGAACTC
rs4346144	ACGTTGGATGATTCACTTTTACAAGACCTC	ACGTTGGATGGCATATTGTACTTAGGAACTC
rs7148078	ACGTTGGATGTGTGTCAGATTGATGGCTTG	ACGTTGGATGCCAAGAGAATAAAGCTGAGAG
rs7148286	ACGTTGGATGGTGGTCATTAAGCTTGCCAG	ACGTTGGATGTGCTATGGATGCTGCTTGAG
rs3783980	ACGTTGGATGTTTTTGCCCCAGGTAAGAC	ACGTTGGATGTGGTGCTTTTGTCTCTCTG
rs1549119	ACGTTGGATGTTTCATCTTCTCTGCCTCC	ACGTTGGATGGTGAAGGCCAGTCATATTGC
rs1984925	ACGTTGGATGAAGTAGCCAGGATTACAGGC	ACGTTGGATGCCAGCCTAGCAAACATGGTG
rs1477261	ACGTTGGATGCAGGGTTATGTGGTATTATC	ACGTTGGATGGGGAAAAGTAAAAGATAAGAG
rs8016447	ACGTTGGATGAATTACAGACGTGTGCCACC	ACGTTGGATGTGACACAGAGAGACTCTGTC
rs7494044	ACGTTGGATGAATTACAGACGTGTGCCACC	ACGTTGGATGTGACACAGAGAGACTCTGTC
rs2023288	ACGTTGGATGGAGAAAAATTGTGATTGATTG	ACGTTGGATGGCCATCAAATCAATCTAATC
rs7151685	ACGTTGGATGACAGTGCTGGCATTACTGGC	ACGTTGGATGTAAAGATCGTCTGCCACTGC

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs2112135	ACGTTGGATGAGTGCAGTGGCCCAATCACA	ACGTTGGATGGTCTAGAGTCCCAGCTACTC
rs2161088	ACGTTGGATGTATAGGGTCTCACTCTTGCC	ACGTTGGATGAGGAGGATCACCTGAGCCTT
rs4903638	ACGTTGGATGATAGGGTGTACTGCGTTGG	ACGTTGGATGAGGCCTAGGTGAGAAGATTG
rs1477262	ACGTTGGATGATGCGTGAGGAGAATGAAGG	ACGTTGGATGAAGGCTAGTGTTCAAGGAAGG
rs1477263	ACGTTGGATGAACCTTCCTGAACACTAGCC	ACGTTGGATGCCTTGCTGCCCCATTTTAAG
rs1477264	ACGTTGGATGCGTAGATAGAACCACCTCAG	ACGTTGGATGAAAGGCGGAGAGCACTTTAC
rs2277917	ACGTTGGATGGCATTGTGTTGCTAGCTGAAG	ACGTTGGATGTTGAACAGGAGTACCGTTTG
rs2277918	ACGTTGGATGTTACGTTCTTACTCAGTCC	ACGTTGGATGACCTGTCGTTTTAAACGCC
rs2277919	ACGTTGGATGTTACGTTCTTACTCAGTCC	ACGTTGGATGACCTGTCGTTTTAAACGCC
rs1978416	ACGTTGGATGAGGGCGTTTAAACGACAGG	ACGTTGGATGCGGGTGAGAGGATATGTTT
rs3759728	ACGTTGGATGATAGTCCCTCGCTGTTTTGG	ACGTTGGATGAGAAAGCACTAGGCCTTTGG
rs6574399	ACGTTGGATGATGCTCTGATGCCATTATGC	ACGTTGGATGAGGGCACGTAACACATCC
rs7155336	ACGTTGGATGGAGGAAGACTCGGTCTAAA	ACGTTGGATGAACAATCTGACACTAGGTGC
rs7156186	ACGTTGGATGATTACGGGTATGAGCCACTG	ACGTTGGATGGAAGTGGACATTAGGTCTGG
rs7142390	ACGTTGGATGTAATCAAGACAGTGTGGTAC	ACGTTGGATGGGGTTTATTTCAAGACTCTC
rs7145875	ACGTTGGATGGTCTTTGAAGCACAAAACC	ACGTTGGATGCTTCATGATCTTGGATTTGGC
rs8014635	ACGTTGGATGGTCTTCTCACTCAAGAACAC	ACGTTGGATGCAACAGAGCAAGACTCCAAC
rs8015938	ACGTTGGATGCCCTGAACTCAAGTGATCTG	ACGTTGGATGATCTAAACAGTGTTCTTGCC
rs8015313	ACGTTGGATGAAACTGATTCTGTACCTGG	ACGTTGGATGGTCAGTCATTTTATAGGCAG
rs8006315	ACGTTGGATGTCACCTGAGGTGAGGAGTTC	ACGTTGGATGCCATGCCTGGCTAAGTTTG
rs6574400	ACGTTGGATGTTATCCTTCTCTGCCAGTG	ACGTTGGATGCCTTTGAACTTCTACCCAG
rs7140816	ACGTTGGATGCAATAGAGTGAGACTCTGTC	ACGTTGGATGATTATGAGCCTCTCTTTAC
rs4566078	ACGTTGGATGTAGAGTCTTGTCTGTACC	ACGTTGGATGAGGAGAATCGCTTGAACCCA
rs7141050	ACGTTGGATGATATGTTATACATATTAGTC	ACGTTGGATGCAAGTTAACCATTATCAACTC
rs3049356	ACGTTGGATGTACCACTGGCAGAGTAGAAG	ACGTTGGATGCACATGGTTTGGGTACTGAG
rs4903639	ACGTTGGATGAGCGAGACTCCGTCTCAAAA	ACGTTGGATGTCAAAGGTAGCCTTGACTGG
rs4903641	ACGTTGGATGACTCCAACCTGGGCAACAGA	ACGTTGGATGCTGGCTCCAGCACACTTATC
rs2364838	ACGTTGGATGTGTAGTCCAGCTACTTGTG	ACGTTGGATGTGATCATAGCTCACTGCAGC
rs2364839	ACGTTGGATGTGGGCAACATAGCAAGATCC	ACGTTGGATGCTCACAAGTAGCTGGGACTA
rs4632066	ACGTTGGATGGAGAAAAAAGAGATGGAGGG	ACGTTGGATGGCCCTGACTGTGTTTTATG
rs2112136	ACGTTGGATGTTTCTTGGGGACTAAGGCTC	ACGTTGGATGTAACAGGCCCTGAAGGAATG
rs4641655	ACGTTGGATGCGATAGAGCAACCCTGTCTC	ACGTTGGATGGCCCTACACCCAGATTCAAG
rs4635269	ACGTTGGATGAAAGTGCTGGGATTACAGGC	ACGTTGGATGCTTGACGATATTTCTGAGG
rs4570764	ACGTTGGATGCTTGACGATATTTCTGAGG	ACGTTGGATGAAAGTGCTGGGATTACAGGC
rs759808	ACGTTGGATGATGAGCTGTGATCATGCCAC	ACGTTGGATGCCTGAACTTCATTGTGCTCC
rs7150531	ACGTTGGATGATGTGCGGTGTGAAGCAAAG	ACGTTGGATGTTGTTTGGCCTGGTCTGATG
rs7154968	ACGTTGGATGTACCCAGGTAACAAACCTGC	ACGTTGGATGTCCCCTATAAGGCTTTCAGG
rs7146657	ACGTTGGATGTGAGTAGCTGGGACTACAGG	ACGTTGGATGTAACACGGTGAAACCCCGTC
rs7145859	ACGTTGGATGAGGCAGGAGAATGGCGTGAA	ACGTTGGATGTTTTGAGACGGAGTCTTGC
rs4903643	ACGTTGGATGTATTCCATGCTGTCTGCCTC	ACGTTGGATGAGTTGACCTTAAAGGCTGGG
rs717682	ACGTTGGATGTTTAGGGACAGAGGCTGAGG	ACGTTGGATGAAGTGAGTGCCCTGATCTC
rs717683	ACGTTGGATGTTGGCAAAAAGGTGGAGGC	ACGTTGGATGTGATGATGGCACAGGGAATG
rs1477259	ACGTTGGATGTGACTGAGACTACCTTCACC	ACGTTGGATGAAGTGCTCACGTAGGTTGTC
rs8019064	ACGTTGGATGCCTTGACAGCAAACTTCAGAG	ACGTTGGATGTGACTGAGACTACCTTCACC
rs8018971	ACGTTGGATGATGGTCTCACTCTGTCACTC	ACGTTGGATGAATTGTTTGAGCCCAGGAGG
rs1477260	ACGTTGGATGAGTGTGATGGTAGCAAGGAC	ACGTTGGATGTGCCATCTGTTTCCCATAGG
rs5809851	ACGTTGGATGACAGAGAGTGTTCAAGCAG	ACGTTGGATGTTGGGCAACAGAGAGAGACT

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs1985149	ACGTTGGATGACTGAAATCTTTGCCTCCCG	ACGTTGGATGGTGGTGCACCTTATGTAGTCC
rs1008988	ACGTTGGATGAGTGTGTCTCAGGGAATGTG	ACGTTGGATGCCTGGCAATTTGTTCTCTGC
rs1008989	ACGTTGGATGGGAATAGCAAGTGTAAACGGC	ACGTTGGATGACTCCAACCGCATCAGCTTC
rs8018222	ACGTTGGATGATCCTCCATATGCTGAACGC	ACGTTGGATGAAGGTGGAACGAGAGACTTG
rs1006040	ACGTTGGATGTTTAGCTCTCTCTCTGTTGC	ACGTTGGATGTCTTGAGCCCAGGAGTTCAA
rs1006039	ACGTTGGATGTGAAGCTGGGAGTTAGAGAC	ACGTTGGATGCCACCATGCCCAGCTAATTT
rs1006038	ACGTTGGATGATAAGCCACTGTGCTCAGTC	ACGTTGGATGGGTAGGGTTTATTAAGTGCC
rs8009784	ACGTTGGATGTGTTTTGGCTATGCTTTGCC	ACGTTGGATGTGACAGAGCGAGACTTTGTC
rs4903644	ACGTTGGATGTTGCAGTGAGCTGAGATTGG	ACGTTGGATGGTGAATGAATGAATAAGGGCC
rs7149496	ACGTTGGATGACAACACACAGTACTGGACC	ACGTTGGATGTGGGTGCATGTTAGAAACGC
rs6574402	ACGTTGGATGCAGGTCCTTTGTCTGACAAG	ACGTTGGATGGGGATGTGCGATTTGATCTG

TABLE 22

dbSNP rs#	Extend Primer	Term Mix
rs7143926	ACCCAAAATTAAGGCAAAATGG	CGT
rs1549071	CACACACATATATACACACACA	ACG
rs8012858	CACACATATATACACACACACA	ACG
rs7155611	GGCACCGTTCTCTCTCTCA	ACT
rs176941	CTGGGCCTCAGTTTACTCAT	CGT
rs176942	AATAGGTTGGTTTGTGCCCC	ACT
rs176943	CCCGTAGTCCCTGTGAAAC	ACT
rs176944	AAAAGTCCACTAATCCTTCCAA	CGT
rs4365221	GAGGGCAACTCAACACATTTTA	ACG
rs3168952	TTGGTGGTGAGATGGACAGA	ACT
rs176945	TACTATACACTCACAAAAATTGTT	ACT
rs176946	TTGTATAACAAAATACCACAAGC	ACT
rs176947	GGCGCCCGCCACTACGC	ACG
rs176948	AAACAGACCTCAGTCTTACA	ACT
rs176949	CTCCCTGCAGTTCCTTGTTA	CGT
rs176950	GGAGTGACATCCCATTACTTT	ACG
rs176951	TCCTCCCCTCCTTGGGTG	ACT
rs7156905	CTGTCTCAAAAAAGGAACCAG	ACT
rs3217197	CTCCAGCCTGAGTGAGAGA	ACT
rs2270443	CAGGTGTGGTGGCTCATGC	ACG
rs176952	CACTGCAATCGCTGCCTCC	ACG
rs176953	GGACCAGCCTGGCCAAACAT	ACT
rs176954	GCAGTGGCTCAATCCCAGC	CGT
rs176955	CTGCCCCCTCCAGCCCTTC	ACT
rs3214416	CATCACGTGTTCTAATGAAAA	CGT
rs176956	AAGCCACCATGCCCAGCC	ACT
rs2544566	CATCTGGGCCTCCCAAAGTA	ACT
rs2544567	CATCTGGGCCTCCCAAAGT	ACT

dbSNP rs#	Extend Primer	Term Mix
rs176957	GGTCTCGATCTCCTGACCT	ACG
rs176958	GGAGTTTTGCTCTTGTTGCC	ACT
rs176959	TTTTATTTGTCCCTTGTTCTTTC	ACT
rs1802227	AATAGTTGCACCAAGCAAGAG	ACT
rs176961	TATGGCAAACCCGTGTCTACA	ACT
rs176962	GGCTCACGCCTGTAATCCTA	ACT
rs7401285	GGGACTCCATATCAGAAAACA	CGT
rs176963	GAAAACACACGCGGGCGC	ACT
rs176964	CAGTGTCTACAAAAGTGCCT	ACG
rs4903631	CTTGAGACAAGATGAAACAGTT	ACG
rs4903632	ATCCTCAGGGAAACGAAAATTA	CGT
rs176965	ATAAGGGCTGCCAGCTTGAT	ACT
rs4903633	TAGCAATTTTATATCTCAGCATG	ACT
rs176966	ACCACACCCAGCTAATTTTTG	ACG
rs176968	TCACACCTGTGACTCCAGC	CGT
rs176969	CCGATTTACTGCATTGCATTTTC	CGT
rs176970	GTACAGTGGGGTGAATAGTTA	ACT
rs7149198	GATATTACTCAGCCATAAAAAAG	ACT
rs7147918	GGATATTACTCAGCCATAAAAAA	ACT
rs7148685	TTGAGACCGTCTATTCAGATC	ACT
rs1184232	GAATGGAAGAAAATGGTTGCAA	CGT
rs1184233	TGCCCAGCCTCTTCAATTAC	ACT
rs1184234	TACCAGCACTTTGGGAGGC	ACG
rs7401998	CCACGCCTGGCTAATTTTTTTT	CGT
rs176974	CTGGGCAACAAAGCAAGACT	ACG
rs6574390	AAATTAGCTGGGTATGATGGC	ACT
rs176975	AATCTAGGTGGTAGGAGATAC	ACT
rs176976	TATAATTCTTTCAGCTTTTCTGTA	ACT
rs176977	TCAGCCTGGGCAACAAGAG	ACG
rs8013727	CCTAACCATAGAAGATAATTAGAA	ACT
rs176978	ATGTTTTGTACCTCTTGTTAC	ACG
rs2111829	GAATTTTGCTTGGTGAACAAAAT	ACT
rs176980	GCTATGAACGCCATTTTATGTA	ACT
rs5809848	TGGGTCTGAAATCCTGCTG	CGT
rs5809849	CGTGTCACTTCCTTTTTTTTTT	ACT
rs4383070	GCCTCCTGAGTAGCTGGG	CGT
rs7493652	TATACCACAAGTAAGTGAATTT	ACG
rs2112133	CCACAACCTGGCTAATTTTTTGT	CGT
rs1963833	CTGGGTGACAGAGCAAGAC	CGT
rs6574391	TGGAGAAGTGATAAACTC	ACG
rs7155062	ATAACCCTTCAAATGAGCATCA	ACT
rs4899674	GGCAAATGGGCTGGGGAG	ACG
rs8022516	AGGCATTCTTTTGGGTATAGTA	ACG
rs7140838	TCTGAAGAATACCAGACCTCT	CGT

dbSNP rs#	Extend Primer	Term Mix
rs7141127	CTGAAGAATACCAGACCTCTC	ACT
rs6574392	CTGGGCACAGCGACTCAC	ACT
rs8003691	TGAAGGCCTCCATGGTATAG	ACT
rs8003979	TCTGAAGCCAGTGAGGAAGT	ACT
rs8010541	GCTAGCTTCAACTCTCCTGAT	ACT
rs8016416	TAACCTTCCTTTGACTTGCTTTTT	CGT
rs8016175	GTCTGCAATCCCGGCACCT	ACG
rs7154571	GTGAGGAGCGTCTCTGCC	ACG
rs7158826	TCGCTCCTCACTTCCCAGA	ACG
rs7159310	CATCTGGGAAGTGAGGAGC	ACT
rs7401900	TGAGAACAGGCCATGATGAC	ACT
rs7160355	CCTGCCAAATCCCCCTCTC	ACG
rs2032781	GAGAAAAGCGGGCAGGACT	ACT
rs6574394	CACCCAGCCTATTTGTATAATT	ACT
rs8007598	CTCTTAACACATTTTTTTACAGCA	ACG
rs2267767	CTGACCTCGTGATCTGCCC	ACT
rs6574395	GGCTCAGGCGATCATCGTA	ACG
rs7150066	CTGCCACCCAAAGTGCTGG	ACT
rs7492334	TTGTGTGTGTGTGTGTGTGG	ACT
rs4359361	GCTTAGTACACTTTAAACATGAT	ACT
rs4605089	TCAGGAACACCGCTTAATTTTT	ACG
rs7146446	CAAGACCTCTTTAAGTAATACTC	ACG
rs4346144	AGACCTCTTTAAGTAATACTCC	ACT
rs7148078	GGCTTGGGTACGGGAAGC	CGT
rs7148286	CATTAAAGCTTGCCAGAAAATCA	ACG
rs3783980	CATCTTCCTCTGCCTCCCA	ACG
rs1549119	CTTCCTCTGCCTCCCATAAAT	CGT
rs1984925	CAGGCACGTGCCACCACA	ACG
rs1477261	AGGAGGAGCCCAAATATGAAA	CGT
rs8016447	CACACCTGGCCATGCTTCC	ACT
rs7494044	CCTGGCCATGCTTCCGTATT	ACG
rs2023288	AATTGTGATTGATTGATTGCGAT	CGT
rs7151685	GTGAGCCACCACATCATCTG	ACT
rs2112135	TCAGGTGATCCTCCTGCCT	ACG
rs2161088	CCAATCACAGCCCACTGCA	ACT
rs4903638	GCCAGAGTGGTCTCCAAC	ACT
rs1477262	AGAGCTCAAGCTGATGTCCT	ACT
rs1477263	TTTTCCTGTTGAGTTCGCATG	ACT
rs1477264	ACCACCTCAGTTTTGCTGTTT	ACG
rs2277917	CCTTGATAACCGCTTGGTCT	ACT
rs2277918	AAAAGCTTCCCGGGGACAG	CGT
rs2277919	AGCTTCCCGGGGACAGCT	ACT
rs1978416	TGAGACTAGCTAATGGAGAGT	ACG
rs3759728	AGCAAATCTACTGCAAACGTG	CGT

dbSNP rs#	Extend Primer	Term Mix
rs6574399	AAGTAGAGCTGCTCCACC	CGT
rs7155336	GAAGACTCGGTCTAAAAAAAAAA	ACT
rs7156186	GCCACTGCACCTGGCCG	ACT
rs7142390	TGGTACTGGCATAAGGATAGA	ACG
rs7145875	CACAAAACCTTAACCTTTTGATTTA	ACT
rs8014635	CAAGAACACTGGTTTTGGTTTT	ACT
rs8015938	CTCAAGTGATCTGCCTGCC	ACT
rs8015313	GATTCTGTACCTGGTTGATCAT	ACT
rs8006315	AACATGGTGAAGCCCCATCT	CGT
rs6574400	GAGATCGCCAGAGACACCA	ACG
rs7140816	TGAGACTCTGTCTCAAATACTA	CGT
rs4566078	CTCAGCTCACTGCAACCTC	ACG
rs7141050	AGCACATAGTAAGTGCCCTAT	ACT
rs3049356	GAATAGTGAAGGTATTGAAATA	ACT
rs4903639	GAGACTCCGTCTCAAAAAAAAAAA	CGT
rs4903641	GGCAACAGAGCGAGACTCC	ACT
rs2364838	CCAGCTACTTGTGAGGCCAA	ACT
rs2364839	AGCCAGACGTGGTGGCAC	ACT
rs4632066	GAGATGGAGGGGGAGCCT	ACT
rs2112136	GGGACTAAGGCTCGCATCC	ACT
rs4641655	GGATTTCTGGGTCCCACTC	ACG
rs4635269	AGCCACCGCGCCCGGCC	ACT
rs4570764	GTGATTATTGGCCGGGCGC	ACT
rs759808	TGCACCACACAGCCTGGG	CGT
rs7150531	AGCAAAGTTAATGGGAGGCC	ACT
rs7154968	AACAAACCTGCATATGTACCC	ACT
rs7146657	CACCCACCACCCCGCCC	ACT
rs7145859	CGGGAGGTGGAGCTTGCA	ACT
rs4903643	GCTCCCTTCTGTCTACTGC	ACT
rs717682	AGGCTGAGGCAGGAGAATC	ACT
rs717683	AAAAGGTGGAGGCCAAAGAC	ACT
rs1477259	CGGAATAATTATATCTGCCTCT	ACT
rs8019064	CAGAGGCAGATATAATTATTCC	ACT
rs8018971	GTCCTCAGGCTGGAGTGC	CGT
rs1477260	GACGAGGAGGAAAGCCATC	ACT
rs5809851	CACAGCAGTGTCTTTTTTTTTTT	ACT
rs1985149	TTCTTCTCCCTCAGCCTCC	ACG
rs1008988	GGGGATGACCTCTCTGGAG	ACT
rs1008989	GCCAGCTTGGCAGATTGAG	CGT
rs8018222	TGCTGAACGCTGGTCCCC	CGT
rs1006040	TGGAGTGCAGTGGCAAGAC	ACG
rs1006039	CATAGCCAGACCCTATGAGA	ACG
rs1006038	ACTGTGCTCAGTCTATGCTG	ACG
rs8009784	ATGCTTTGCCTTAAAGTGGTG	ACT

dbSNP rs#	Extend Primer	Term Mix
rs4903644	GCCTGGGCAACAGAGCAAG	ACT
rs7149496	GATTCTGTAAGTCTGGTATGAG	ACT
rs6574402	CTGACAAGAAAATGACTGCATA	ACT

Genetic Analysis

[0242] Allelotyping results are shown for cases and controls in Table 23. The allele frequency for the A2 allele is noted in the fifth and sixth columns for osteoarthritis case pools and control pools, respectively, where “AF” is allele frequency. The allele frequency for the A1 allele can be easily calculated by subtracting the A2 allele frequency from 1 (A1 AF = 1-A2 AF). For example, the SNP rs7143926 has the following case and control allele frequencies: case A1 (A) = 0.75; case A2 (T) = 0.25; control A1 (A) = 0.71; and control A2 (T) = 0.29, where the nucleotide is provided in paranthesis. Some SNPs are labeled “untyped” because of failed assays.

TABLE 23

dbSNP rs#	Position in SEQ ID NO: 3	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs7143926	218	76161268	A/T	0.25	0.29	0.216
rs1549071	1440	76162490	C/T	0.15	0.20	0.098
rs8012858	1442	76162492	C/T	0.93	0.95	0.335
rs7155611	2611	76163661	C/T	0.02	0.02	0.949
rs176941	4317	76165367	A/C	0.31	0.35	0.271
rs176942	4724	76165774	A/G	0.02	0.02	0.911
rs176943	4788	76165838	G/T	0.13	0.18	0.037
rs176944	5202	76166252	G/T	0.09	0.14	0.107
rs4365221	5780	76166830	C/T			
rs3168952	5974	76167024	C/T			
rs176945	6644	76167694	C/G	0.95	0.96	0.801
rs176946	7430	76168480	A/G	0.10	0.15	0.054
rs176947	7938	76168988	C/T	0.10	0.08	0.473
rs176948	8095	76169145	C/T	0.31	0.35	0.132
rs176949	8183	76169233	A/C	0.03	0.02	0.887
rs176950	8312	76169362	C/T	0.78	0.70	0.008
rs176951	8352	76169402	A/C			
rs7156905	9348	76170398	C/T	0.89	0.90	0.794
rs3217197	9378	76170428	-/TCTC	0.29	0.35	0.036
rs2270443	9617	76170667	A/G	0.39	0.34	0.176
rs176952	9727	76170777	C/T	0.17	0.24	0.018
rs176953	9834	76170884	C/T			
rs176954	9899	76170949	G/T	0.43	0.52	0.010
rs176955	10211	76171261	C/T	0.12	0.18	0.028
rs3214416	10377	76171427	-/T	0.91	0.89	0.544
rs176956	10695	76171745	C/T	0.51	0.49	0.492
rs2544566	10729	76171779	C/G			
rs2544567	10730	76171780	C/T			
rs176957	11433	76172483	A/G			
rs176958	11951	76173001	C/G	0.02	NA	NA

dbSNP rs#	Position in SEQ ID NO: 3	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs176959	12697	76173747	C/T	0.30	0.34	0.147
rs1802227	12982	76174032	A/C	0.92	0.95	0.332
rs176961	14419	76175469	C/T	0.51	0.47	0.158
rs176962	14501	76175551	C/T	0.82	0.79	0.192
rs7401285	14983	76176033	A/C			
rs176963	15280	76176330	C/T	0.51	0.46	0.155
rs176964	15475	76176525	A/G	0.53	0.49	0.197
rs4903631	15888	76176938	A/G			
rs4903632	15976	76177026	A/T			
rs176965	16307	76177357	A/C	0.55	0.52	0.368
rs4903633	16442	76177492	A/C	0.83	0.83	0.970
rs176966	17255	76178305	C/T			
rs176968	18948	76179998	G/T	0.23	0.27	0.246
rs176969	19435	76180485	A/T	0.14	0.20	0.052
rs176970	19753	76180803	C/T	0.35	0.38	0.328
rs7149198	20021	76181071	C/T			
rs7147918	20022	76181072	A/C			
rs7148685	20503	76181553	A/G	0.19	0.18	0.669
rs1184232	20590	76181640	G/T	0.16	0.19	0.316
rs1184233	21804	76182854	G/T	0.36	0.36	0.895
rs1184234	21919	76182969	C/T	0.36	0.35	0.797
rs7401998	21990	76183040	A/T			
rs176974	22412	76183462	A/G			
rs6574390	22536	76183586	C/T			
rs176975	23432	76184482	A/G	0.18	0.23	0.147
rs176976	23468	76184518	G/T	0.86	0.80	0.087
rs176977	23772	76184822	C/T	0.42	0.41	0.794
rs8013727	24325	76185375	C/T			
rs176978	24773	76185823	C/T	0.10	0.12	0.512
rs2111829	26274	76187324	C/T	0.02	NA	
rs176980	27440	76188490	C/G	0.79	0.73	0.018
rs5809848	28561	76189611	-/ACAG	0.11	0.16	0.091
rs5809849	30071	76191121	-/A	0.60	0.57	0.355
rs4383070	31764	76192814	A/T			
rs7493652	33008	76194058	C/T			
rs2112133	35310	76196360	A/T			
rs1963833	35460	76196510	A/C			
rs6574391	37112	76198162	A/G	0.69	0.63	0.064
rs7155062	37285	76198335	A/G	0.17	0.18	0.878
rs4899674	37747	76198797	C/T	0.57	0.52	0.201
rs8022516	38057	76199107	C/T	0.57	0.51	0.135
rs7140838	38859	76199909	A/C	0.17	0.17	0.957
rs7141127	38860	76199910	A/G			
rs6574392	39525	76200575	A/G	0.27	0.32	0.099
rs8003691	40216	76201266	A/G	0.70	0.63	0.029
rs8003979	40281	76201331	C/T	0.10	0.15	0.024
rs8010541	41453	76202503	C/G	0.38	0.38	0.993
rs8016416	42091	76203141	A/T	0.09	0.14	0.035
rs8016175	42513	76203563	A/G			
rs7154571	42935	76203985	C/T			
rs7158826	42985	76204035	A/G			
rs7159310	43003	76204053	A/G	0.62	NA	
rs7401900	43281	76204331	A/G			
rs7160355	43716	76204766	C/T			
rs2032781	43866	76204916	A/G	0.80	0.74	0.047
rs6574394	44234	76205284	G/T	0.61	0.54	0.091
rs8007598	44596	76205646	A/G	0.09	0.10	0.734

dbSNP rs#	Position in SEQ ID NO: 3	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs2267767	44871	76205921	C/T			
rs6574395	45005	76206055	A/G	0.10	0.14	0.203
rs7150066	45282	76206332	A/C	0.91	NA	
rs7492334	47178	76208228	A/C			
rs4359361	47816	76208866	G/T			
rs4605089	47887	76208937	A/G			
rs7146446	48134	76209184	C/T	0.09	0.09	0.981
rs4346144	48135	76209185	A/G	0.83	0.85	0.368
rs7148078	48276	76209326	G/T	0.44	0.50	0.098
rs7148286	48400	76209450	C/T	0.96	0.96	0.893
rs3783980	48798	76209848	A/G	0.15	0.20	0.073
rs1549119	48803	76209853	A/T	0.18	0.25	0.027
rs1984925	49146	76210196	C/T	0.04	0.04	0.882
rs1477261	49969	76211019	A/T			
rs8016447	51059	76212109	A/G	0.10	0.15	0.049
rs7494044	51064	76212114	C/T			
rs2023288	53285	76214335	A/T	0.97	0.98	0.774
rs7151685	54560	76215610	C/T			
rs2112135	54748	76215798	A/G	0.05	NA	
rs2161088	54785	76215835	C/G			
rs4903638	55102	76216152	C/G	0.59	0.59	0.975
rs1477262	55644	76216694	A/G	0.12	0.17	0.040
rs1477263	55705	76216755	G/T	0.18	0.23	0.057
rs1477264	55841	76216891	A/G	0.45	0.42	0.271
rs2277917	56623	76217673	C/G	0.30	0.36	0.039
rs2277918	56825	76217875	A/C	0.49	0.45	0.232
rs2277919	56827	76217877	A/G	0.20	0.17	0.310
rs1978416	56892	76217942	C/T	0.79	0.73	0.074
rs3759728	59150	76220200	A/T	0.13	0.18	0.083
rs6574399	59958	76221008	A/T	0.33	0.36	0.396
rs7155336	60231	76221281	C/T	0.25	0.28	0.250
rs7156186	60524	76221574	A/G	0.85	0.85	0.965
rs7142390	61871	76222921	C/T			
rs7145875	62226	76223276	C/T			
rs8014635	63230	76224280	G/T	0.07	0.11	0.062
rs8015938	63468	76224518	G/T	0.08	0.07	0.693
rs8015313	63787	76224837	C/T	0.67	0.71	0.135
rs8006315	65732	76226782	A/C			
rs6574400	65989	76227039	A/G	0.75	0.70	0.099
rs7140816	68832	76229882	G/T	0.54	0.48	0.095
rs4566078	69904	76230954	C/T			
rs7141050	70365	76231415	A/G			
rs3049356	70886	76231936	-TATC	0.64	0.69	0.091
rs4903639	73088	76234138	A/T			
rs4903641	73103	76234153	C/T	0.54	0.66	~0.0001
rs2364838	75934	76236984	C/T			
rs2364839	75966	76237016	C/T	0.18	0.18	0.988
rs4632066	76273	76237323	C/T	0.66	0.66	0.961
rs2112136	77943	76238993	C/T	0.70	0.64	0.064
rs4641655	78466	76239516	C/T	0.52	0.48	0.174
rs4635269	78861	76239911	C/T			
rs4570764	78872	76239922	A/G	0.55	0.68	~0.0001
rs759808	79836	76240886	G/T	0.12	0.18	0.043
rs7150531	80908	76241958	C/T	0.33	0.31	0.491
rs7154968	81509	76242559	C/G	0.03	NA	
rs7146657	83576	76244626	C/T	0.57	NA	NA
rs7145859	83662	76244712	C/G			

dbSNP rs#	Position in SEQ ID NO: 3	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs4903643	83782	76244832	C/T	0.10	0.14	0.074
rs717682	84282	76245332	G/T	0.11	0.13	0.624
rs717683	84444	76245494	A/G	0.79	0.75	0.121
rs1477259	85129	76246179	C/G	0.11	0.16	0.022
rs8019064	85151	76246201	A/G	0.90	0.93	0.192
rs8018971	85296	76246346	A/C			
rs1477260	85809	76246859	C/G	0.12	0.16	0.085
rs5809851	86387	76247437	-/T	0.30	0.30	0.993
rs1985149	86494	76247544	A/G	0.22	0.23	0.892
rs1008988	89786	76250836	A/G	0.61	0.58	0.380
rs1008989	89894	76250944	A/T	0.14	0.18	0.172
rs8018222	90122	76251172	G/T			
rs1006040	92067	76253117	A/G	0.13	0.18	0.092
rs1006039	92187	76253237	C/T	0.06	0.10	0.133
rs1006038	92312	76253362	A/G	0.19	0.24	0.114
rs8009784	92824	76253874	G/T	0.13	0.18	0.037
rs4903644	93733	76254783	C/T	0.41	0.38	0.383
rs7149496	96553	76257603	C/G			
rs6574402	96941	76257991	A/C	0.12	0.17	0.037

[0243] Allelotyping results were considered particularly significant with a calculated p-value of less than or equal to 0.05 for allelotype results. These values are indicated in bold. The allelotyping p-values were plotted in Figure 1C for the discovery cohort. The position of each SNP on the chromosome is presented on the x-axis. The y-axis gives the negative logarithm (base 10) of the p-value comparing the estimated allele in the case group to that of the control group. The minor allele frequency of the control group for each SNP designated by an X or other symbol on the graphs in Figure 1C can be determined by consulting Table 23. For example, the left-most X on the left graph is at position 76161268. By proceeding down the Table from top to bottom and across the graphs from left to right the allele frequency associated with each symbol shown can be determined.

[0244] To aid the interpretation, multiple lines have been added to the graph. The broken horizontal lines are drawn at two common significance levels, 0.05 and 0.01. The vertical broken lines are drawn every 20kb to assist in the interpretation of distances between SNPs. Two other lines are drawn to expose linear trends in the association of SNPs to the disease. The light gray line (or generally bottom-most curve) is a nonlinear smoother through the data points on the graph using a local polynomial regression method (W.S. Cleveland, E. Grosse and W.M. Shyu (1992) Local regression models. Chapter 8 of Statistical Models in S eds J.M. Chambers and T.J. Hastie, Wadsworth & Brooks/Cole.). The black line provides a local test for excess statistical significance to identify regions of association. This was created by use of a 10kb sliding window with 1kb step sizes. Within each window, a chi-square goodness of fit test was applied to compare the proportion of SNPs that were significant at a test wise

level of 0.01, to the proportion that would be expected by chance alone (0.05 for the methods used here). Resulting p-values that were less than 10^{-8} were truncated at that value.

[0245] Finally, the exons and introns of the genes in the covered region are plotted below each graph at the appropriate chromosomal positions. The gene boundary is indicated by the broken horizontal line. The exon positions are shown as thick, unbroken bars. An arrow is placed at the 3' end of each gene to show the direction of transcription.

Example 7

In Vitro Production of Target Polypeptides

[0246] cDNA is cloned into a pIVEX 2.3-MCS vector (Roche Biochem) using a directional cloning method. A cDNA insert is prepared using PCR with forward and reverse primers having 5' restriction site tags (in frame) and 5-6 additional nucleotides in addition to 3' gene-specific portions, the latter of which is typically about twenty to about twenty-five base pairs in length. A Sal I restriction site is introduced by the forward primer and a Sma I restriction site is introduced by the reverse primer. The ends of PCR products are cut with the corresponding restriction enzymes (*i.e.*, Sal I and Sma I) and the products are gel-purified. The pIVEX 2.3-MCS vector is linearized using the same restriction enzymes, and the fragment with the correct sized fragment is isolated by gel-purification. Purified PCR product is ligated into the linearized pIVEX 2.3-MCS vector and *E. coli* cells transformed for plasmid amplification. The newly constructed expression vector is verified by restriction mapping and used for protein production.

[0247] *E. coli* lysate is reconstituted with 0.25 ml of Reconstitution Buffer, the Reaction Mix is reconstituted with 0.8 ml of Reconstitution Buffer; the Feeding Mix is reconstituted with 10.5 ml of Reconstitution Buffer; and the Energy Mix is reconstituted with 0.6 ml of Reconstitution Buffer. 0.5 ml of the Energy Mix was added to the Feeding Mix to obtain the Feeding Solution. 0.75 ml of Reaction Mix, 50 μ l of Energy Mix, and 10 μ g of the template DNA is added to the *E. coli* lysate.

[0248] Using the reaction device (Roche Biochem), 1 ml of the Reaction Solution is loaded into the reaction compartment. The reaction device is turned upside-down and 10 ml of the Feeding Solution is loaded into the feeding compartment. All lids are closed and the reaction device is loaded into the RTS500 instrument. The instrument is run at 30°C for 24 hours with a stir bar speed of 150 rpm. The pIVEX 2.3 MCS vector includes a nucleotide sequence that encodes six consecutive histidine amino acids on the C-terminal end of the target polypeptide for the purpose of protein purification. Target polypeptide is purified by contacting the contents of reaction device with resin modified with Ni^{2+} ions. Target polypeptide is eluted from the resin with a solution containing free Ni^{2+} ions.

Example 8

Cellular Production of Target Polypeptides

[0249] Nucleic acids are cloned into DNA plasmids having phage recombination sites and target polypeptides are expressed therefrom in a variety of host cells. Alpha phage genomic DNA contains short sequences known as attP sites, and *E. coli* genomic DNA contains unique, short sequences known as attB sites. These regions share homology, allowing for integration of phage DNA into *E. coli* via directional, site-specific recombination using the phage protein Int and the *E. coli* protein IHF. Integration produces two new att sites, L and R, which flank the inserted prophage DNA. Phage excision from *E. coli* genomic DNA can also be accomplished using these two proteins with the addition of a second phage protein, Xis. DNA vectors have been produced where the integration/excision process is modified to allow for the directional integration or excision of a target DNA fragment into a backbone vector in a rapid *in vitro* reaction (Gateway™ Technology (Invitrogen, Inc.)).

[0250] A first step is to transfer the nucleic acid insert into a shuttle vector that contains attL sites surrounding the negative selection gene, ccdB (*e.g.* pENTER vector, Invitrogen, Inc.). This transfer process is accomplished by digesting the nucleic acid from a DNA vector used for sequencing, and to ligate it into the multicloning site of the shuttle vector, which will place it between the two attL sites while removing the negative selection gene ccdB. A second method is to amplify the nucleic acid by the polymerase chain reaction (PCR) with primers containing attB sites. The amplified fragment then is integrated into the shuttle vector using Int and IHF. A third method is to utilize a topoisomerase-mediated process, in which the nucleic acid is amplified via PCR using gene-specific primers with the 5' upstream primer containing an additional CACC sequence (*e.g.*, TOPO® expression kit (Invitrogen, Inc.)). In conjunction with Topoisomerase I, the PCR amplified fragment can be cloned into the shuttle vector via the attL sites in the correct orientation.

[0251] Once the nucleic acid is transferred into the shuttle vector, it can be cloned into an expression vector having attR sites. Several vectors containing attR sites for expression of target polypeptide as a native polypeptide, N-fusion polypeptide, and C-fusion polypeptides are commercially available (*e.g.*, pDEST (Invitrogen, Inc.)), and any vector can be converted into an expression vector for receiving a nucleic acid from the shuttle vector by introducing an insert having an attR site flanked by an antibiotic resistant gene for selection using the standard methods described above. Transfer of the nucleic acid from the shuttle vector is accomplished by directional recombination using Int, IHF, and Xis (LR clonase). Then the desired sequence can be transferred to an expression vector by carrying out a one hour incubation at room temperature with Int, IHF, and Xis, a ten minute incubation at 37°C with proteinase K, transforming bacteria and allowing expression for one hour, and then plating on selective media. Generally, 90% cloning efficiency is achieved by this method. Examples of expression vectors

are pDEST 14 bacterial expression vector with att7 promoter, pDEST 15 bacterial expression vector with a T7 promoter and a N-terminal GST tag, pDEST 17 bacterial vector with a T7 promoter and a N-terminal polyhistidine affinity tag, and pDEST 12.2 mammalian expression vector with a CMV promoter and neo resistance gene. These expression vectors or others like them are transformed or transfected into cells for expression of the target polypeptide or polypeptide variants. These expression vectors are often transfected, for example, into murine-transformed adipocyte cell line 3T3-L1, (ATCC), human embryonic kidney cell line 293, and rat cardiomyocyte cell line H9C2.

[0252] Modifications may be made to the foregoing without departing from the basic aspects of the invention. Although the invention has been described in substantial detail with reference to one or more specific embodiments, those of skill in the art will recognize that changes may be made to the embodiments specifically disclosed in this application, yet these modifications and improvements are within the scope and spirit of the invention, as set forth in the claims which follow. All publications or patent documents cited in this specification are incorporated herein by reference as if each such publication or document was specifically and individually indicated to be incorporated herein by reference.

[0253] Citation of the above publications or documents is not intended as an admission that any of the foregoing is pertinent prior art, nor does it constitute any admission as to the contents or date of these publications or documents. U.S. patents and other publications referenced herein are hereby incorporated by reference.

Nucleotide and Amino Acid Sequence Examples

[0254] Table A includes information pertaining to the incident polymorphic variant associated with osteoarthritis identified herein. Public information pertaining to the polymorphism and the genomic sequence that includes the polymorphism are indicated. The genomic sequences identified in Table A may be accessed at the http address www.ncbi.nih.gov/entrez/query.fcgi, for example, by using the publicly available SNP reference number (*e.g.*, rs552). The chromosome position refers to the position of the SNP within NCBI's Genome Build 34, which may be accessed at the following http address: www.ncbi.nlm.nih.gov/mapview/map_search.cgi?chr=hum_chr.inf&query=. The "Contig Position" provided in Table A corresponds to a nucleotide position set forth in the contig sequence (see "Contig Accession No."), and designates the polymorphic site corresponding to the SNP reference number. The sequence containing the polymorphisms also may be referenced by the "Nucleotide Accession No." set forth in Table A. The "Sequence Identification" corresponds to cDNA sequence that encodes associated target polypeptides (*e.g.*, Q96FX2). The position of the SNP within the cDNA sequence is provided in the "Sequence Position" column of Table A. If the SNP falls within an exon, the corresponding amino acid position (and amino acid change, if applicable) is provided as well. Also, the allelic variation at the

polymorphic site and the allelic variant identified as associated with osteoarthritis is specified in Table A. All nucleotide and polypeptide sequences referenced and accessed by the parameters set forth in Table A are incorporated herein by reference. Genomic nucleotide sequences for *KIAA0296*, *chrom 4* and *SNW1* regions are set forth in SEQ ID NO: 1-3, respectively. A polymorphism in Table A designated by "AA" is present in the genomic nucleotide sequence of SEQ ID NO: 4, which follows Table A.

Table A

RS_ID	Chromosome	Chrom Position	Contig Accession No. [1]	Contig Position	Nucleotide Accession No. [2]	Sequence Position	Amino Acid Position	Locus	Locus ID	A [3]	Allelic Variability	OA Assoc. Allele
552	3	16276963	Hs3_22673_34:16	16241963	XM_209584	intergenic		Q96FX2	253631; 92106	R	[A/G]	A
12904	1	152323489	Hs1_79549_34:1	1556529	NM_004428 NM_018845	mma-utr locus-region		EFNA1 LOC5597 4	1942 55974	F	[A/G]	A
2282146	20	49881706	Hs20_11519_34:8	14249192	NM_002827	coding-synon	303	PTPN1	5770	R	[C/T]	T
734784	20	44409056	Hs20_11519_34:8	8776542	NM_002251	coding-nonsynon	V489I	KCNS1	3787	R	[G/A]	G
1042164	19	13125398	Hs19_11452_34:1	4527200	NM_004907	coding-nonsynon	V133A	ETR101	9592	R	[T/C]	T
749670	16	31124685	Hs16_24968_34:1	2488292	NM_014699	coding-nonsynon	G327E	KIAA029 6	9726	F	[C/T]	T
955592	2	85570358	Hs2_22340_34:13	64411756	NM_032213	coding-nonsynon	I70T	FLJ21977	84173	R	[T/C]	C
1143016	1	2209625	Hs1_4507_34:16	317874	NM_002617	coding-synon	93	PEX10	5192	R	[T/C]	T
755248	1	39408404	Hs1_33153_34:6	1601557	NM_024732	mma-utr		FLJ14351	79787	R	[G/A]	G
1055055	1	42293810	Hs1_33153_34:6	4486963	NM_173642	coding-synon	308	MGC478 16	284716	F	[A/G]	A
835409	1	59157267	Hs1_33153_34:6	21350420	NM_018291	mma-utr		FLJ10986	55277	F	[T/G]	G
927663	1	155318902	Hs1_79549_34:1	4551942	AL138899	10739		LOC2540 79	254079	F	[T/G]	G
8162	2	42552684	Hs2_22340_34:13	21394082	NM_004718	mma-utr		COX7A2 L	9167	F	[A/G]	G
831038	2	170404402	Hs2_5560_34:14	20411275	NM_004525	intron		LRP2	4036	F	[C/T]	T
33079	2	231577157	Hs2_5560_34:14	81584030	NM_003113	intron		SP100	6672	R	[G/A]	A
1710980	3	11027339	Hs3_22673_34:16	10992339	AC024910.5	71095		SLC6A1	6529	F	[C/A]	A
4472029	3	102614466	Hs3_5769_34:14	7788711	NM_020357	intron		PCNP	57092	F	[T/A]	A
799570	3	113007639	Hs3_5769_34:14	18181884	NM_145753	intron		LL5beta	90102	F	[A/G]	G
1282730	3	113026933	Hs3_5769_34:14	18201178	NM_018394	intron		FLJ11342	55347	R	[G/A]	A
1518875	3	188430421	Hs3_5769_34:14	93604666	none	intergenic		none	none	R	[T/C]	C
1568694	4	8584835	Hs4_6464_34:15	505708	NM_003501	intron		ACOX3	8310	R	[G/A]	A

RS_ID	Chromosome	Chrom Position	Contig Accession No. [1]	Contig Position	Nucleotide Accession No. [2]	Sequence Position	Amino Acid Position	Locus	Locus ID	A [3]	Allelic Variability	OA Assoc. Allele
905042	4	16229175	Hs4_6473_34:15	7153650	none	intergenic		none	none	R	[A/T]	A
1957723	4	36919171	Hs4_16453_34:15	4170634	none	intergenic		none	none	R	[G/A]	A
794018	4	40066604	Hs4_16453_34:15	7318067	none	intergenic		none	none	R	[G/A]	G
707723	4	47757791	Hs4_6395_34:10	7544215	NM_006587	UTR/intergenic		PRSC	10699	F	[C/T]	T
893861	4	47872633	Hs4_6395_34:10	7659057	NM_000087	mma-utr		CNGA1	1259	R	[G/A]	G
1914903	4	125442695	Hs4_16510_34:16	49479005	none	locus-region		LOC377201	377201	R	[G/A]	A
2062232	4	163485648	Hs4_16762_34:16	24456077	NM_020116	intron		DKFZp566D234	56884	F	[C/T]	T
26609	5	53447781	Hs5_6588_34:13	3996839	NM_019087	intron		FLJ20051	54622	R	[A/T]	T
1370987	5	79365937	Hs5_6870_34:13	8673996	XM_293971	locus-region		LOC345778	345778	F	[A/G]	A
1012414	5	150535958	Hs5_29448_34:10	11670384	NM_001155	intron		ANXA6	309	R	[G/A]	G
435903	5	156713884	Hs5_23289_34:11	1542547	NM_014376	intron		CYFIP2	26999	R	[G/A]	G
1248	5	156742512	Hs5_23289_34:11	1571175	NM_014376	intron		CYFIP2	26999	F	[T/A]	T
703508	5	169797195	Hs5_23289_34:11	14625858	NM_004137	mma-utr		KCNMB1	3779	R	[G/A]	A
226465	6	1755892	Hs6_35042_34:3	1750892	NM_001500	intron		GMDS	2762	R	[G/C]	C
241448	06	32785099	NT_033951	4177541	NM_000544	UTR		TAP2	6891	F	[C/T]	C
763155	6	35240628	Hs6_7749_34:13	25990900	NM_152753	intron		CEGF3	222663	R	[A/C]	A
1040461	6	57102190	Hs6_7749_34:13	47852462	NM_016277	coding-nonsynon	S207G	RAB23	51715	R	[T/C]	C
462832	6	111731861	Hs6_25897_34:13	15855720	NM_002912	intron		REV3L	5980	R	[A/T]	A
804194	6	112053555	Hs6_25897_34:13	16177414	NM_002037	intron		FYN	2534	R	[T/C]	C
1022646	6	112074901	Hs6_25897_34:13	16198760	NM_002037	intron		FYN	2534	F	[A/G]	A
1569112	7	699341	Hs7_7976_34:14	250677	NM_006869	intron		CENTA1	11033	R	[G/A]	G
805623	10	29478204	Hs10_8862_34:15	11449523	none	intergenic		none	none	F	[A/G]	G
1019850	11	83985464	Hs11_34082_34:6	14458185	NM_001364	intron		DLG2	1740	R	[A/T]	A
1599931	12	56545117	Hs12_29578_34:1	20402156	none	intergenic		none	none	F	[A/G]	G
AA	13	29711866			NM_006644	exonic		HSP105B	10808	F	[A/G]	G
279941	13	101396169	Hs13_10109_34:1	16787844	NM_000452	mma-utr		SLC10A2	6555	F	[T/G]	G

RS_ID	Chromosome	Chrom Position	Contig Accession No. [1]	Contig Position	Nucleotide Accession No. [2]	Sequence Position	Amino Acid Position	Locus	Locus ID	A [3]	Allelic Variability	OA Assoc. Allele
1062230	14	23479277	Hs14_26604_34:1	5409277	NM_014178	mma-utr		STXBP6	29091	F	[C/T]	C
1859911	14	70531426	Hs14_26604_34:1	52461426	NM_004296	intron		RGS6	9628	R	[T/C]	T
1477261	14	76211019	Hs14_26604_34:1	58141019	NM_012245	intron		SNW1	22938	F	[T/A]	T
1191119	14	98332328	Hs14_26604_34:1	80262328	NM_004434	intron		EML1	2009	R	[G/A]	G
657780	15	60862497	Hs15_10351_34:1	33936998	none	intergenic		none	none	F	[A/G]	A
1393890	15	95036282	Hs15_10431_34:1	12272041	XM_096908	UTR		LOC145945	145945	R	[G/C]	C
1478714	16	7132995	Hs16_10709_34:1	4474136	NM_018723	intron		A2BP1	54715	R	[G/A]	G
868213	16	66996694	Hs16_10655_34:1	15942010	NM_178516	intron		LOC283849	283849	F	[C/T]	C
690115	17	73471192	Hs17_10798_34:1	6766230	NM_178128	mma-utr		LOC283985	283985	R	[G/A]	G
1465501	18	18066783	Hs18_11123_34:1	1303898	none	intergenic		none	none	F	[A/G]	G
899173	19	14094734	Hs19_11452_34:1	5496536	NM_018154	intron		ASF1B	55723	F	[C/T]	C
10477	20	38236897	Hs20_11519_34:8	2604383	NM_015568	mma-utr		PPP1R16B	26051	F	[C/T]	T
926393	20	38331457	Hs20_11519_34:8	2698943	NM_021931	intron		DHX35	60625	F	[C/T]	C
465271	21	26507612	Hs21_11669_34:9	13247612	none	intergenic		none	none	F	[C/T]	C
13847	21	44259022	Hs21_11672_34:1	715558	NM_020132	locus-region		AGPAT3	56894	F	[A/G]	G
738658	22	30168847	Hs22_11677_34:9	11234862	NM_019843	intron		EIF4ENIF1	56478	F	[C/A]	C

[1] Contig Accession Number which can be found in the NCBI Database:
http address: www.ncbi.nih.gov/entrez/query.fcgi

[2] Sequence Identification or Nucleotide Accession Number which can be found in the NCBI Database:
http address: www.ncbi.nih.gov/entrez/query.fcgi

[3] "A" column is the sequence orientation ("F" is forward, "R" is reverse).

AA genomic sequence (SEQ ID NO: 4)

TCATTAGCTTTTTTCAGTTTTTCACATTCCTGATACAGACGTAGGAGTGCTCGTATTTT
GGATTTTGCATCCAACTTGTACTTAGTTTTAAATTCTGCACA[A/G]AAATGTTCCACT
AACTTTTCATCGAAGTTTTTTCCTCCTAAGAAAGGATCAAAAGCTGTTCCCAGTACC
TAATTTGGTTGAAACAACAAATAGTCTGGTT

[0255] Following are genomic nucleotide sequences for a *KIAA0296* region (SEQ ID NO: 1), a *chrom 4* region (SEQ ID NO: 2), and a *SNW1* region (SEQ ID NO: 3). The following nucleotide representations are used throughout: "A" or "a" is adenosine, adenine, or adenylic acid; "C" or "c" is cytidine, cytosine, or cytidylic acid; "G" or "g" is guanosine, guanine, or guanylic acid; "T" or "t" is thymidine, thymine, or thymidylic acid; and "I" or "i" is inosine, hypoxanthine, or inosinic acid. Exons are indicated in italicized lower case type, introns are depicted in normal text lower case type, and polymorphic sites are depicted in bold upper case type. SNPs are designated by the following convention: "R" represents A or G, "M" represents A or C; "W" represents A or T; "Y" represents C or T; "S" represents C or G; "K" represents G or T; "V" represents A, C or G; "H" represents A, C, or T; "D" represents A, G, or T; "B" represents C, G, or T; and "N" represents A, G, C, or T.

KIAA0296 genomic sequence (SEQ ID NO: 1)

>16:31076951-31174000

1	ccccaccccc	caacagctgc	acagctctgga	gcgaatatac	acgcccacca	cccacacacc
61	caagacccaa	tacacttttt	taaactttat	ttttacttct	atztatttat	ttttaattat
121	tttttaaaaa	tctaattaga	gatgaggtct	taggctgggc	acagtggctc	atgcctgtaa
181	ccccagcact	tcgggaggcc	gaggcaggca	gatacacgag	cgaggagatc	acgaggtcag
241	gagttcRaga	ccagcctggc	caatatgggtg	aaaccccatc	tctgctaaaa	atacaaaaat
301	gagctgggcg	cggtggtgtg	cacctgtaat	ctcagctact	tgggaggctg	aggcagaatt
361	gtttgaaetc	aggaggcgga	tgctgcagtg	agctgagatc	gtgccactgc	actccagtct
421	gggagacaga	gcgagactac	gtctcaaaac	aaacaaacaa	acaaacaaca	acaacaaaaa
481	cagagataag	gtcttggcat	gtgtcccagg	ctggtctcaa	gtcctgggct	caaaggattc
541	tcctgcctca	gcctcccaaa	gtgctaggat	tacaggcgtg	aaccactgca	cccaccctac
601	tttttttttt	tttttttttt	atacaggatc	tcactctgtc	acccgggctg	gagtgcagtg
661	gcaagatcac	tgctgactgt	acccttgacc	tcaggggactc	aagtgatcct	cctgcctcag
721	cctcctgagt	agctgggact	acaggagagc	gccagcacac	ctgggtaatt	aagatttttt
781	ttgtagagac	agacgctatg	ttgccaggc	tgctctcgaa	ctcctggcct	caagtatac
841	acccttggcc	tcctaaagtg	ctgggatcac	aggcatgagc	cactgcacct	agcctaatat
901	agttaatatc	cccgtaagg	ctgctcagag	ggcctgagag	gaacaaagg	ctcagctctg
961	gagagctcca	ccccagcgc	caatctctct	aaatggcctc	tttcctctcc	atattccacc
1021	acaaggcttg	gagtccagct	tcctgtgacc	ttaagtcacc	attccaaagc	cctgcgatct
1081	caccagaga	ccacaagtga	aataatatta	taatcctgag	aagtttagtg	gaccaagatg
1141	gcatgccatc	aagacgctga	gaaacaaaga	ggaagatggg	accagggggc	ccagaagacg
1201	ctggaaccca	cagtattaaa	agctcagaga	ggctgggcac	agtggctcac	acctgtaatc
1261	ccagcacttt	gggaggccaa	ggtgggtgga	tcacttgagc	ccaggggttt	gagaacagcc
1321	tgggcaacat	ggcgaaaccc	agtctctacc	aaaaaatata	caaaaattag	ccaggcatgg
1381	tggtgcgtgc	cttagtacca	gctacttggg	aggctgaggc	aggaggattg	actgaacctg
1441	agagcacacc	actgcactcc	agcctggatg	acagaaccag	acctgacctc	aaagagaaga
1501	aaaaaaaaaa	aaaaaaaaag	cccagagggg	agggYaccct	caacagtttt	ccagcccctt
1561	ccacatcctt	cctaacctca	cttgatagtg	ttcaagtcct	accttaggca	aggcagaaat
1621	tataggacca	agccgcctaa	tggggaaatt	gagtcaccaga	gagaagtaat	gcattattta
1681	agatcccatg	caggactatg	agtcaggggg	ccaagagccc	ttccaccgtg	tgccactcag

1741	agacacagag	taggaggggg	aaggggggctg	gggtggcagg	gacaaaagat	gcaggaggca
1801	agcagcagtg	actgaagagg	cagagggtga	catgaaagac	ccaggagcag	agaatctttc
1861	cttatcatct	ccaggggaca	ccactgggca	gggcttggcc	tccggaaaaa	ccctgcattc
1921	cctctgtggg	ttcatcaggg	caccactctc	ctactagctg	ggtttttttt	ttttgttttg
1981	ttttgttttt	gagacagagt	cttactctgt	cacctaggct	ggagtgaat	ggcgtgatct
2041	cagctcactg	taacctccac	ctcccatggt	caagcaatcc	tcctgtctca	gcctcccaag
2101	tagctgggat	tacaggcacc	tgccatcatg	cctggctaata	ttttgtatgt	ttgtagagac
2161	agggtttcgc	catgttggcc	aggctgggtc	ccaactcctg	gcctcagggt	atctgcctgc
2221	ctcagcctcc	caaagtgtct	ggattacagg	catgagccac	cacaccctgc	ctgagctggg
2281	ttttaacagg	aagaggagaa	gagccaaaac	tcctcacata	gaatcacaca	gcacttgaca
2341	gtttccaacc	tcatcatcac	tgaagttag	agcagccgat	acccaYaaag	atgatctccc
2401	catcccccta	cagttaccca	ctgtgcagag	ggagatccac	acttagagac	aggaagcgat
2461	ttccagaagt	ccatagcaac	tcagtccag	gaatctaggt	ttcctgacca	gggcatagca
2521	gaaagggtcc	attccttttc	ttgcttgtag	cttcacagaa	gcttcctgga	cagagccctg
2581	gggtccagga	gacctgttat	tcattcccgg	ctatgctgag	acttgctgag	tgaccttggg
2641	gactccttct	agagaatata	agtcccacga	gtgcaggaa	ttttgtctat	tagtccttga
2701	tgtatctcca	gccctagaac	agtgtttggc	ccatactatg	tgcccaaaaa	atatccatta
2761	aatgactgaa	tggtgctgtg	catggtgggt	catgcctgta	atcccagcac	tttggaagc
2821	tgaggcagaa	ggattgctta	agcccaggag	ttagagacca	gcttggacaa	catagtgaga
2881	ccgcactctg	taaaaatttt	taaaaataaa	aaatgagtga	atatctagat	agccaggatt
2941	agagaagtgt	cacagtcaga	aagcctgaag	cctaaagaag	accaaggaa	caggggcttt
3001	atcctcagat	acatgaaagc	ctgaaattct	gtccacaagt	atttatagag	ggcccgtaat
3061	gttcttggta	ctgggctagg	aactccccag	attcagttaa	gaacaaagtc	attacctggc
3121	ctcagatgca	aggcaggggc	tgggggggtg	gagtggcagg	gaggcagcgt	gatcaataca
3181	aacacttttc	ttagcctgag	ctgccctgac	atggctctgac	ggctcacaag	gtgggtgagt
3241	cagccgggct	gcagtgttca	aggagggcgc	cggctggcgc	cccacctgtc	agagcgtcgc
3301	ccagaaggat	gcggaagaag	agattttctgc	cttggctgag	gtcacttccc	acccccagat
3361	tccttgccca	cacaaccctg	caattttctg	acgctgacga	ctcggatcct	attatttccc
3421	gattttcaag	gtcccattgat	gctgacagcc	ccaaatgcta	agtcgtcagt	ccgcccacgc
3481	cctggaccgc	aaagcaataa	aggcgagggtc	agcaagggtc	ctaccaccca	ctgcctcgaa
3541	aggcctctgg	gggtggctcg	cgcgcctctc	cccacctcgc	ggggggcctg	tggcgctcgc
3601	tcggctcgtt	gggtgcccgg	gacgtcgtga	tgagaacggc	gtcccagaga	cggcggtgac
3661	agagccggga	cacgtgacag	tcacaggggtc	acattctgcg	gtccacgagt	ttgggaccgg
3721	gctggtcacg	tgacgcgggt	ggggcaccat	ggggtgatgt	gagatgcggg	tgtctcggat
3781	tacgtacaaa	tgacgtattc	ctaccccttt	tggcaaccag	atttccgttg	gaagatgcaa
3841	cggttccggt	gacggtagca	agttctcgcg	tccaggcata	tcgcgttccg	ctcgggctgc
3901	aacaacttcc	gactccacct	tcccagcctc	gggcaaggaa	gagacgcgac	catgtgcgca
3961	tgccccgaat	ttatcacgga	ggggcggggc	tgaggctgcg	ggagctggag	cggggaagaa
4021	aagggaattc	caacctgttg	aaccttgggg	ggtccccggg	gtcggcgccct	tcccattgac
4081	tgtgggcggg	gcaagggacg	gagcctctgg	cggctcgtgg	gggtgttggg	gtccgcaggg
4141	ggagggaggg	gagtgtcaga	gtgtgtagcg	ggtacgggaa	ttccaaattt	gagggcctcc
4201	cggctctggc	gccggggagg	gagagctcag	gccgccatgc	gggacaggac	ccacgagctg
4261	agacaggtga	gacgccaggg	cagcggggat	ggggacgggc	ggacgaaactg	gaacgcagga
4321	cttctggtct	tcgggatagg	gaggggtggc	tgatggccag	gaaggaaagt	cccggaaagc
4381	tgtgggtcct	gcggggtaag	agccgcagcg	aaacgggtgt	gccaatgact	ccgggcctgg
4441	cagggggatg	acagctcgga	cgaagaggac	aaggagcggg	tcgcgctggt	ggtgcacccg
4501	ggcacggcac	ggctggggag	cccggacgag	gagttcttcc	acaaggtaag	gggctggggg
4561	ctccgcctgg	attcgcgagg	gtgtaggagg	acccgaggag	tagcgtgggt	tggagtaccc
4621	catatctctt	tcagccctct	cggtcaccct	ccccagggtc	ggacaattcg	gcagactatt
4681	gtcaaactgg	ggaataaagt	ccaggagtgt	gagaaacagc	aggtcaccat	cctggccacg
4741	ccccttcccc	aggagagtga	gtgaaacccc	ggctgcaggg	cgcatgctcc	gccccaggga
4801	ttgtgggggt	tgtagttcca	cgcagggtgt	ggccagagtg	gtttgttgag	gtgggggctg
4861	ctgtttggga	gtcttggcct	tctcttattc	aggcatgaag	caggagctgc	agaacctgcg
4921	cgatgagatc	aaacagctgg	ggagggagat	ccgcctgcag	ctgaagggtg	agctcctggg
4981	acctcagaca	gatccttccc	tctgatcctg	ccctgttgtt	ggtatatctg	gggagtgtgt
5041	ggcccagaga	agccagtgat	atatccaggt	cacacagcag	gcctgggtct	agcatctgtc
5101	tcctggcctc	caggccattg	tactctccac	agcacaagtc	cgcctctcag	gttcttttat
5161	ttacaatgaa	accattttact	tacacagtta	tcgctgcccc	ctgggcattc	tttgggcagg
5221	gagatggagt	tttgttaggt	ggcctctgca	tacctatggg	aactcagtga	tgtaattcaa
5281	agaaaaataa	acttactttc	tctctctaga	ggctcagcct	tagtcatttt	atgataaatt
5341	atatttccct	aaaaatccta	tggagacaag	tacccccaat	acccctgtgt	cttcccacag
5401	ccatagagcc	ccagaaggag	gaagctgatg	agaactataa	ctccgtcaac	acaagaatga
5461	gaaaaaccca	ggtgggtttt	ttttctcaga	aatgaggaca	tttcagcaaa	tgtttcatga
5521	agtattagat	gacagggtga	tgaagggaag	gcctgcagag	atcatggagt	ccaattggat

```

5581 gactttttccaa aatgggggaaa ctgagctcag agagagaaaag aacttgctca aggtcaggaa
5641 gccagggtctc ctgatgtctca gtccgggttat aacaccctgc tttattttct tccattcaat
5701 aggaaggttac tgtgacccca gacaagacct agtcttggct gtgggacaca tgttttcttt
5761 tctttttttg cctcagcctc ctgaatagct gggattacag gcggacaccc ccatgccag
5821 ctaatttttg tagttttagt agagactggg cttcaccatg ttggccaggc tggtttcgaa
5881 ctccctgacct caggtgacct tcctgctcg gectcccaaa gtgctgaatt acaggcgtga
5941 gccaccatgc ccagctggga cacatgtttt ctgggagtca agatgaggag ttagggttca
6001 ataggggata aagacattac tcacgtggga cctggtggct aacggcgctg cccagggaag
6061 gagagtgaga agtcataaat gactggcagg tttcctatct atgtgacagg gacatcctta
6121 gtcccacagg tggaattcaa gaagttagga agaggaaactt ccttggggca aactgaaga
6181 ggaactcccc tgggtgtgata tcttattttt ttaattatta ttattttttt ttattttttt
6241 ctactctgt cctcagggtt ggagtacagt ggcacaatct tggctcactg caacctccac
6301 ctccctcaag cgattctcct gcctcagcct cacgagtagc tgggattaca ggtgtgcacc
6361 accacacctg gctaattttt tatatttttg gtagaaatga ggtttcacct tgttggccag
6421 gctggtctcg aactcctgaY ctcaactgat ccactgctt tggcctcgca aagtgtatg
6481 attataggca tgagccaccg cgcgcggccc ctggggtgat atcttagtaa ggaggttg
6541 agtgatctga ctggccctct ctgggtcccc agtgaggagg ataccaggag gtcagggttg
6601 gagtagtttg gccagggtt cagcagggac cccagattga agatggagca gcttgggcat
6661 cttggaaggg tgaagctgga accaggaaaag cagatgtatc tctggaaaag gaactccaag
6721 gaatgagcat atttaaggcc tcagaagaag gggcaaggca gagcagatgc cccagaacca
6781 gtgtttctgg ggaagcctgt ggtgtgattt ggcattgagt gttgagggtc catgtgggcc
6841 tgttgacact gtttcgcccc ggcaacatgt tcatctctag gcgtaggagc tgtgtgttag
6901 gcagcgaggt tggcattcag caagcattca gcagttacat attgggtgcc tactgtgtgc
6961 cagacctttt tggaaactgt taggatacag cagtgaacca gtgatccctg tctcatgga
7021 acttcccttc tgggtgtgac aatcaccata ataaataagt gaattattta gaacataata
7081 agcattaagg aaaaaagagc aggggaagag ggactaagca tgctggagga ggtgagttg
7141 cagttgaaag caggtggagg aagcttcatt cagaaggtaa catctgaaca agagacttaa
7201 aggtgtttgc tgggaatgag cattctaggt agaaggaaaa gtgaatgcaa aggtttaagc
7261 tgagagtgtg ctttgtctag ggagggttaa ggagaccagt gtggatgggc agaggaaggg
7321 aacagtaaga ggaagtaaga tcagagaggt catgggagaa ggagagatca tagagggcta
7381 gccaggcacc gttgctcacg cctgtaatcc cagcactttg gaggctgagg tgggaggtt
7441 ggttgagccc aggagtttga gaccagcctg ggcaatatag tgagaccccc cccctttttt
7501 tttttttcct ttgagacagg gtctcactct gttgcccagg ctggagtaca gtggtgccat
7561 ctctgtctac tgcaacctcc gcctcctggg ttcaagccat tctctgtctt cagcctcca
7621 agtagctggg actacaggcg cccaccactg caccaagcta atttctgtac ttttagtaga
7681 gatggggttt caccacgttg gccaggctgg ccttgacctc ctgacctcag gtgatccacc
7741 tgctcagacc tcccaaagtg ctgggatcac aggcattgagc caccgtgccc ggccaacctt
7801 gtctctatta aaaataaaaa taggccagggt gcagtggctc acgcctgtaa tggaggccga
7861 ggcagggtga tcacaagggtc aagagatcaa gaccatcctg gccaacatgg tgaacccca
7921 tctctactaa aaatacaaaa attagcgtg cgtggtggcg cgtgcctgta gtcccagcta
7981 ctcgaggagg tgaggcaaga gaattgtctg aaccggggag gccaaagggtg cagtgaaggc
8041 agattgtgccc actgcactcc agcctgggca acaagagtga aactctgtct caaaaaacaa
8101 ataataaata aataaataaa taaataaata aataaataaa taaaaaagat catggaggac
8161 cacataggcc tgataagggc tttggttttt agtctaagag aaatggggga gcctgtcaag
8221 gtcatacaaa ggtggttaag gtggcagatc ccgcataaga gctcatgcta tttgtcact
8281 gtactatggg gttgccgagg caccgacggg gcagggatcc tcccaggggc actcagctta
8341 tattcttcat ctttagcatg gggctcctgtc ccagcaattc gtggagctca tcaacaagtg
8401 caattcaatg cagtccgaat accggggagaa gaacgtggag cggattcgga ggcagctgaa
8461 gatcagttag ttgtgcatgc ccagcctggc ccgagggggc aggtaatccc aaccaacccc
8521 tgagcctggc cttttccttc acagccaatg ctgggatggt gtctgatgag gagtggagc
8581 agatgctgga cagtgggcaa agcggaggtg ttgtgtccaa tgtgagtggc cacagccagc
8641 cctctctgct tgtgctctcc atccccctct agtctgtctc gtttctcgac ctctgggct
8701 cagggtgatcc tctgctcctc gcctcccag tagctgggac tatagggtga agccactgca
8761 ccccgcttgc tgtggccctt tctgattaag ggcacctga ggcctctaag ggaattaatt
8821 agctgcctg gagtcacca tcagattcca ggctgagggc tcccagaag ctcaacagga
8881 gtttctgacc gtgtgtcggt ctccctgtga acagtggccc cactcctgtc cccccccag
8941 atcctgaagg acacgcaggt gactcgacag gccttaaatg agatctcggc ccggcacagt
9001 gagatccagc agcttgaacg cagtattcgt gagctgcacg acatattcac ttttctggct
9061 accgaagtgg agatgcaggt ggggtgcccc cgagcccca gacgtgagac caggctcagt
9121 ccaaactgcc agMctcccgc caYcctttaga ttctctccct gaggcttttg tgtcttcag
9181 gtttgcccat gccccagat tgggtgcttat tctatcctt agctgtaccc cgagaatggc
9241 acctgcctct gctgctacac agatgcccac tccctctctg atagcaccct gccccctctc
9301 caaaacttga gcctgcccag gtctggcccc agcctcact cccctccac taacagcatc
9361 cacccttata cctctcagag gtccagtcag agttgcccta gaggggctgc ctctaacat

```

9421	ctgtacaagg	ctgggggtggg	ggcgggcgttc	ccctggccct	ggttgtgagt	Wgagttgagc
9481	ttccagccct	gtcctggagg	agctggcctc	agtcattgcta	cagccaatgc	ccttttgagc
9541	ctgagactta	caggaaagag	atctcattca	gtaggagtac	tgagacctga	ggctgggtgg
9601	gccaggagga	ggcagggata	gggagggctt	tgcagcagct	gtagataggc	ctggaagaat
9661	gggtaaattc	agacagattt	gtgaaggcac	agttcaccat	ctgtgaaagg	tatgagccat
9721	ttgagccct	tagctccaag	ctaccactgc	agatagaggt	tgtatgggat	aagtgagcag
9781	gggacaagg	actacatgat	agaaggggcc	tggaaagccat	ccccaaggag	tctgaacttt
9841	tgtcagatca	agtcttgccc	ttgtctttgt	tagtgcaatt	tttttttccct	gccaggaatg
9901	ttcttcagtc	atctgggggtg	gggtgggcaa	aggcatcctt	acctccctga	accaccccat
9961	cctctgagca	gggggagatg	atcaatcgga	ttgagaagaa	catectgagc	tcagcggact
10021	acgtggaacg	tgggcaggag	cacgtcaaga	cggccctgga	gaaccagaag	aagcgaggga
10081	aggtgagcct	ccaggcccg	gccactgccc	caggcaccct	gtgtgacttc	cctgaccccc
10141	tcctctccca	cagaagaaag	tcttgattgc	catctgtgtg	tccatcaccg	tcgtcctcct
10201	agcagtcac	attggcgcta	cagtgggttg	ataatgtcgc	acattgttgg	tgagatgttg
10261	tgggctgccc	cctggcctgc	cccagccctg	gccccagccc	tccctcctcc	ctcagacctt
10321	gttctccctc	cttctcctac	aggcactagg	agcaccagga	acccagggcc	tggccttctc
10381	tcaccagcag	ctgggggggca	gggcagagcc	tccagtcgga	ccccctcctc	acactggccc
10441	ctatgcagaa	gggcagacag	ttcttctggg	gttggcagct	gctcattcat	gatggcctcc
10501	tccttcaggc	ctcaatgcct	gggggagggc	tgcactgtcc	tgattggccg	ggacacacgg
10561	ttttgtaaaa	aattaaaaaa	caaaaaaaga	gcatagaaa	ccctgtgcac	gtgtgttctc
10621	ggaagggtg	ggccaaggct	tccgggcata	caacctcctt	acctcctgga	cgcccgaggc
10681	ggcaggctg	ggcctggctg	ctcagtgcaa	actgccaggg	gtgctgtgcc	cacagcaggc
10741	tgggttctgcc	tttctgcacc	cccataggaa	tgggtgggca	gggaggggta	acaccggcat
10801	ctagctcctg	gctcagtact	gtccccggga	aaggaccact	gtgagtatct	gtcttggaag
10861	tgatgaggct	gaccaggcca	ggctgggacg	cagggtgagat	gggggtttgg	gtggcatcag
10921	tgggccttct	tgtggcccag	aggaagaggg	accatgaaaa	aatgcctaata	tgagctgtgc
10981	actttggatg	cagtggatag	ggatgctctg	gtttcagcag	ggatgacatt	ggagtgggat
11041	gttaagctgg	ggaagagggt	gccagtcaga	aagcacagga	ggccggggccc	tgtgaccaac
11101	aaaagcatca	tctttttacat	aagcgttttag	gcagggtgtg	gtggctcgca	ccagtaatcc
11161	cagcactttg	ggaggcccag	gcaggaggat	ctcttgagcc	caggagcttg	agaccagcct
11221	aggcaggatg	gggcaacctc	ttctcttttag	agaataataa	ttttacaaat	tgaccaggcg
11281	tgatggcaag	tgtctgtcgt	cccagctact	ccagaggctg	agggtggagg	atcgcttgag
11341	cccaggagat	taaggctgca	gtgagccatg	gtcatccac	tgactccat	cctgggtgac
11401	agagcgagac	cctgtctcaa	aaataatagc	aatcatcatc	agtagcagca	gcagcagcag
11461	cagcagcata	gagagccagt	gatcctggat	cagtgcacct	ggttgctgag	ggttacctgg
11521	ctgaagcagg	tgttggcagc	agaaaagcct	gacctctgat	ttcttccata	aggtaacctga
11581	aatccaagcc	ctgactaaat	ttcttttttt	cttttttttt	gagacagagt	cttgttcttt
11641	tgcccaggct	ggagtgcagt	ggcactatct	cagctcactg	caagctccgt	ctcccagggt
11701	cacgccattc	tctctgcctca	gcctcccag	tagctgggac	tgcaggcacc	cgccaccaca
11761	cccggtctaat	tttttgtatt	tttagtagag	acgggggttt	accgtgttag	gatggctctg
11821	atctactgac	atcgtgatct	gcctcctcgc	gcctcccaaa	gtgttgggat	tacaggcggt
11881	agccaccgcc	taaatttctta	agggtcctta	gtcctgatgc	ctaatttctg	gagtggacgt
11941	ggctcctggt	cccgcacacc	tagagttttt	gtttgtttgt	ttgtttgttt	tgagacagag
12001	tctcgtctctg	tcgcccagcc	tgggtgagc	tggcgcaatc	tcggctcact	gcaagctccg
12061	cctcccgggt	tcacgccatt	ctcctgcctc	agcctccaga	gtagctggga	ctacaggcgc
12121	ccgccaccat	gcccggctaa	tttttttttt	ttttttgaga	cggagtcttg	ctctatcgcc
12181	cagactggag	tgcagtgggtg	cgatctccgc	tcactgcaaa	cttcgcctcc	cgggttcacg
12241	ccattctcct	gcctcaggct	cctgagtagc	tgggactaca	ggcaccggcc	accgcgccc
12301	gctaattttt	tgtattttta	gtagagacgg	ggtttcatcg	tgttagccag	gatggctctg
12361	atctcctgac	ctcgtgatcc	gcccgcctca	gcttcccaaa	gtgctgggat	tacaggcatg
12421	agccaccgag	cccggccccc	gacacctagt	tttaaagggt	aagccggctc	ctggcacctg
12481	cctacttgca	gtaggggcgc	gcctagctct	gacctccaa	gtctggggac	tgctgcagc
12541	ccgcccagtc	catcccactt	tcaatcttac	aggcccctgc	tgttgctgcc	gctgccgcgc
12601	ctcccagctg	cccagctctg	cgggctcagt	cccgcgttgc	catgtgtggg	agaccgcgtc
12661	gcgtaagcgc	tggatgtggc	ttcgctgag	cacattggac	cgggctctgg	actgggctag
12721	gggaagggca	ggagggcgga	attggggccc	agggccaggc	ctcgccgacc	cccgactcgc
12781	cctcccgggtg	gccccgcagc	gcctcccggg	ggccctggag	tgcaggtctt	accgtccgag
12841	atcgtccgca	actgggcgag	ctgtgcatgg	ggcgtggcta	aggccgtggg	ttggttacga
12901	ttggccagcg	ggacttaagt	gttgtctctg	aagagcatgg	acattagtct	ggagggtcct
12961	gggaagtgga	tccccgcccc	accatcaaat	ggcgcttagg	tctaggaagc	gggtgtgggt
13021	gggaccttag	ggcgaggcgc	agacataccc	cgaagtgggt	ggattgtata	ccgcaagggg
13081	ctggatcgaa	ccccccaaag	acactggaag	gctgtgtggc	tgaggagggc	ccggcagatc
13141	cagtgtgtcS	tgggctttac	aggaaagagc	tccaccttct	ctggagtgtg	cagatgcgat
13201	ctaggtgtgt	ccacccgatg	ggagctgcgg	gcccggcaga	tgctgcccc	gtacaaagct

13261	gatttgacc	tggggcctct	ggacttccct	gatttctctgc	ttgcatctcc	agcaaagtcc
13321	tgccccgttg	gctgccttca	tccactctct	cacttctctg	ccttcagagt	aaaattgcaa
13381	gatctgtggt	gcttactggg	atctgataga	gtctctcggc	atccactgtc	tatgcagcgg
13441	gtgtccacct	gcagcggggg	ccatgtgcag	cggggggcca	cgtgcagtgt	gtgcctcttc
13501	ttagccatgc	tggacagcgc	cgccctgaa	aagcagctcc	cgggtttcac	ccagaaagcc
13561	atccagaacc	tcctggaana	ggtggcctga	tggccaagtg	gcctcggatg	ccaggtcaa
13621	tcctttgaac	ttttcctgtg	ggctgtcagg	acccatagaa	ggtctttgag	caggtgagtt
13681	tggagcagat	ctggtaggca	agcgaacaga	tggatgYgtg	cactggagat	tccgtgggtt
13741	cccctgtgta	catctcttcc	ctttgggaaa	ctgccctgag	tgaggggcta	agggcaggat
13801	ttgcattgaa	atcctagctt	tgctgctgtc	agcccaactt	ttaggcaaca	gggtcttggg
13861	ttgatgtgac	atttccaagt	ccatcttgta	tcacaacctg	tcagctgcag	ctcacttatt
13921	caatctattg	tggttcaagt	tcccaagaaa	atgaatcagt	ctgggtctgt	ctccagatca
13981	gattacgttt	acttgccatg	gaattgtctg	ccctttaact	caagactttg	cactgtttgt
14041	cacatttgta	atcccagcac	tttgggaggc	caaagcagga	gtattgcttg	agcccaggag
14101	ttcaagacca	gccagggaaa	tataacaaga	ccctatctct	acaaaaatta	aaattaggtt
14161	gggactgtg	gtcatgcct	gtaatcccag	cactttggga	ggccctggca	ggtggatcac
14221	ctgatgtcag	gcgttcgaga	ccagcctgac	caacatggtg	aaaccccgct	tctactaaat
14281	acaaaaagtt	agctggatat	ggtggtgcag	gcctgtaatc	ctacttggga	ggctgaggca
14341	gaagaatcac	ttgaacccgg	gaggtggagg	ttgcagtgcag	ccgagattgt	cccattgcac
14401	tccaacttgg	gcaacaagag	caaaactccc	tctcaaaaaa	aaaaaaaaaa	aaaaaaaaagc
14461	caggtRcatg	tcagtggtag	gtgctgtggt	tccagctac	ttgggaggct	gaggtgggag
14521	gattgcttgg	gcctgggggt	gagaccacag	tgagccaata	ttgcaccact	gcactccagc
14581	ctggacaaca	gaataatacc	ctgtctcaaa	aaaaaaaaaa	aaaaaaaaaga	aaaaaaaaagaa
14641	aagaaaaaga	ctttgccctt	gagtcagagc	tttacccttt	tacccttggc	taagatggat
14701	gtaggaagtg	acatggtaca	aaatgctgca	gcagagcgtg	tgtatgtgct	ggaagaggag
14761	ttgactaggg	cagtgattga	catctctgtt	ccagatattt	gcttaccttc	cctgtgggc
14821	ccctccctat	aggagcatta	tatgtctcatt	ccctacttac	aatagggttg	gctataggac
14881	ttgctttggc	cagtgggaata	tgggtaggaa	ggcaaaatat	cggccgggcg	caatggctca
14941	cacctgtaat	cccagcactt	tgggaggatg	aggcgggtgg	atcagctgag	gtcaagagtt
15001	cgagaccagt	ctggccaacg	tggtagaacc	ctgtctctac	taaaagtaca	aaattagcca
15061	ggcatgggtg	cacgggcctg	taatcccagc	tgctaggaag	gctgaggcag	gagaaactc
15121	tgaacccggg	aaggaggagg	ttgcaatgag	cccagatcat	gccattgcac	tccagcctgg
15181	acaaaaagtg	aaactccgtc	tcaaaaaaaa	aaaaaaaaag	taaagtatca	ctctgcata
15241	gaagctttag	ggcaccattg	agtactctag	cagcttccag	tctcttctcc	ctctgctcag
15301	gctcataaac	ctggcagttt	ccagatctag	acttctcttt	cagcctgcaa	cccagaatga
15361	caatgacatg	aagctgggct	acagcctacc	tataaaatga	tgcaagaatt	aagaaataaa
15421	tctctcttgc	tgtgagccat	tgatatatgg	agggtgtttg	ttagcacatc	caaatgttta
15481	aacaaactgt	tacagaatta	ataccagaa	gtggtgtgct	gcaacaataa	aaattgagcc
15541	tcagccgggc	acgggtggctc	acacctgtaa	tcccagcatt	ttgggaggcc	aaggtaagtg
15601	ggtcacctaa	ggttaggagt	ttgaaaccag	cctggccaac	atgacaaaac	cctgtctcta
15661	ctaaaaatac	aaaaaaaatt	agccaaggat	ggtggtagg	gctgtaatc	ccagctagag
15721	gctgaggcag	gagaatcgct	tggacccagg	aggcagaggt	tggcagtMg	tcaagattgc
15781	gccactgcac	tccagcctgg	gcgatggagt	gagactccat	ctcaaaaaat	taaaaataaa
15841	aaataaaaa	attattaaaa	attagccagg	tgtgatggca	tgtgcctgta	gtcccagcta
15901	cttgggaggc	tgagatggga	ggatcacttg	agcccaggaa	gcagaggttg	cagtgcacca
15961	agattgcacc	actgcactcc	agcctgggtc	caaaaaaaa	aaaatcccca	gccaggcatg
16021	gtggctcatg	cctgtaatcc	cagcactttg	ggaggctgag	gtgggtgctg	aggctcaggag
16081	tttgatacta	gcctggcaaa	catggtgaaa	ccctgtctcc	actaaaaata	caaaaattat
16141	ccaggcatgg	tgggtgggcac	ctgtaatccc	agctactcag	gaggctgagg	ctggagaatc
16201	ggttgaacct	tggatgcgga	ggttgcaagt	agccaagatc	aagccactgc	actccagcct
16261	ggcgacacaga	gcaagactat	ctcaaaaaaa	aaaaaaaaaa	cctaaactat	gtaaacata
16321	tgacattgac	ggtgagctgg	acagtggctg	gtaagggaac	tgtcattgga	agttggaaag
16381	atggtgacgt	gtgttatgca	atggtgaatc	gtttggttaa	actgtaagct	tatgacaaa
16441	tgagctttag	gcttttaggt	aagaactggg	gaaagggagt	attggtagca	tgctgtcact
16501	actattgcat	gcatttgagg	agttactaga	agaaaagagt	gactcagaaa	ttaattggct
16561	agtttataag	cagaaatgga	agagaaatata	gaaattcgag	gcaagtgtac	cacattttca
16621	gtaaaagata	caactgagaa	agtccttgag	ccacaagggt	ttcgtttttg	tttttgagac
16681	agtcttgctc	ttgtttccaa	ggccaccttc	tgggttcaag	cctttctcct	gactcagcct
16741	cccaagtagc	tgggattaca	ggcgtgcacc	accacgtcca	gctaattttt	gtattttcag
16801	tagagacagg	tttcacatg	ttggccaagg	tggcttgaa	cttctgacct	caaattgatcc
16861	tcccacctcR	gcctcccaaa	gtgctgggat	tacaggtgtg	agccactgcg	accggctgag
16921	ctacaagttt	tgattaaaag	tcactcttgt	ggcaaggggc	atatcaagta	tatggctatt
16981	atgccctttg	taaaaatctc	caaaactgatc	aaagtgttcc	ctaataaatc	ctctcagcta
17041	gtcaagatga	ttcaaaggaa	agagggttaa	agtgtaaact	accttggctg	ggcgtgggtg

17101	ctcacgcctg	taattccagc	actttgggag	gctgaggtgg	gcggatcacc	tgaggtcagg
17161	agtttgagac	tagcctaacc	aacatggaga	aaccccgctc	ctactaaaaa	tacaaaatta
17221	gccaggcatg	gtggtgcatg	cctgtaatcc	caactacttg	ggaggtcag	gcaggagaat
17281	tgcttgaacc	tgggaggcgg	aggttgcagt	gagccaagat	cacccatggc	actccagcct
17341	gggcaacaag	agtgaaaactc	catctcaaaa	aaaaaaaaaa	atgtagctta	cctgagggag
17401	tcagtaggct	caactacagt	taagtctaac	gtcatggta	tgtctgaaaa	gaattatggg
17461	tatgctgttg	acccatggat	ctgaatggag	taaaatacgt	aagttcagtt	ttggagggaa
17521	ttgccctgct	tccccctgct	aacacccccct	caccctgaca	aaaagccacc	aggttaaatc
17581	ttgaccatga	gtgttcaata	cttagtatga	tttttaggtc	cccaagtttc	ttcttttttt
17641	tttattttcg	agaccgggtc	tcactctgtc	accagcctg	gagtgcagt	atgcaaccac
17701	agctctctat	aacctcgaac	ttctgggctc	acacgatcct	cctgcctcag	cctcccaagt
17761	agctgggact	acaggcccat	gccacccag	caggctaatt	tttgtttttc	aaattttttt
17821	gaaacaaaat	ctcactctgc	caccaggtc	gaagtgcagt	ggcacgatct	tggtcactg
17881	caacctccgc	ttcttgggct	tgagtgtacc	acttacctca	gcctcccaag	tagctgggac
17941	tacaggtgtg	cgctaccatg	cccggcta	ttttgtattt	ttttggtaga	gacagggctc
18001	tgctatattg	cccaggtcgg	tctcgaaact	ctgaactcaa	gcgattcacc	tgtcttggcc
18061	tcccaaagtg	ctggcattat	aggcgtgcag	tgtaccacca	tgcccagcct	atttttgttt
18121	tgttttgctt	tggtttgttt	tgagatgaag	tcttgctctg	tcactccagc	tgaggtgcag
18181	tggcacaatc	aagcctcact	gcagcctcta	cctctagggc	tccagtgtat	ccccacctc
18241	agccttctga	gtagctggga	ctacaggtc	gcgccaccac	acctggctaa	tttttctatt
18301	tttttctgga	gaggttttca	gcctgttgcc	caagctggct	ttgaacttct	ggctcttaag
18361	agtttctcct	gcttggcttc	ccaaagtgtat	gggattacag	gtgtgagcca	ccatgccag
18421	cctaattttt	gtatttcagg	tttttttttg	ttttgttttg	ttttgttttt	agtagagatg
18481	ggggtctctg	tatgttgccc	aggctggcct	caagcaatcc	ttgcctcaag	tgatcctcct
18541	gcctYagcct	ctcaaaatac	tgtgattgca	gatgtgaacc	accatgcccg	gcctgggtct
18601	ccaaatttct	tttttttttt	tagagacgga	gtctcgttct	gtcaccaggg	ctggagtgc
18661	gtggtgtgat	ctcggctcac	tgcaagctct	gcctcccagg	ttcacgccgt	tctcccgcct
18721	cagcctcccg	agtagctggg	actacaggYg	cccgccacca	tgcccggcta	attttttttt
18781	gtatttttag	tagagacagg	gtttcactgt	gttcgccagg	atgggtttcg	tctcctgacc
18841	tcgtgatctg	cccgcctcgg	cctcccaaa	tggtgggggt	ataggcgtga	gccaccgcac
18901	ctggccatgg	gtctccaaat	tctatgggc	atgaaggaga	ctgagaaaag	ttctactctt
18961	cagaaagaca	taaccaccag	tgtcctctca	attgtggcca	aggagaataa	gtggaaaagg
19021	gtggtttact	ctaagggcag	agccaagaac	atgggtgaaga	atgaactagg	gaactcttcc
19081	cactcccagg	gaaaagtggg	ggttcttctc	aacatctgcc	caKcagcact	ttagacttag
19141	tggggcccag	agcctgctgt	gtgtctcctg	tccttccttc	cttttttttt	tttttttttt
19201	tttttgagac	agagtttcac	ttttgtcacc	catgtttag	tgcaatggca	ctatctcgcc
19261	tcactgcaac	ctctgcctcc	tgggttcaag	cgattctctt	gcctcagcct	cccagtagtc
19321	tgggactaca	ggtgcatgcc	accacgcctg	gctaattttt	ggttttgggg	ttttttgttt
19381	ttgtttttga	gacggagtct	tgcactgtcg	cccaggctgg	agtggaaatg	cacgatctcg
19441	gtcactgca	acctctgcct	cctgggttca	agcgattctc	ctgtctcagc	ctcctgagta
19501	gtgggacta	caggggccc	ccaccagccc	cggctaactt	tttgtatttt	tagtagagac
19561	cgggtttcac	tatgttggcc	tggctgtgtc	tgaactcctg	acctgtgtat	ctgccccctt
19621	cggcctccca	aagtgtctga	attacagacg	tgagccactg	cgctgggcta	atttttgtat
19681	tttttagtaga	gacaggtttt	caccgtgttg	gccagggtgg	tctcaaactc	ctgacctcag
19741	gtgatccacc	ggcctcggcc	tcccaaagtg	ctgggattac	aggcatgagc	caccgcaccg
19801	ggccctgtcc	ttccttctga	acgggagtgt	gctctgctgt	tctcctgtcc	ttgttctgct
19861	ttatatgttg	gatgtgttgc	tgtgtgtgtg	tgtgtagaaa	tggggcacag	gtaacttgtc
19921	tctgtctctc	tcttattttg	tagctcatag	gtctctgaat	caagagaagc	cacatctgga
19981	cctgatatag	aagagactat	tagagatcct	gggcttgagg	ctgattccat	gtcagatggg
20041	tcacttaggt	ggtctccctt	gggaagggga	tgcatttatt	ttgcataatg	aagaaaatgc
20101	aaaggcagta	tttgtaagga	agagggcaga	cgggggaaga	ttttataatt	gttcaaaaac
20161	attcactggg	atgtgtgtgg	tggctcacgc	ctataatccc	agtgcttttg	gaggggtgaag
20221	caggaggatc	acttgaggcc	aggagtttga	gaccagcatg	ggcaacatag	tgagacccta
20281	tctctacaaa	aaataaaaaca	ataaaaaaaa	attagctggg	cgtggtggtg	cttgccctgta
20341	gtcctagcta	cttaggaggc	tgaggtggga	ggatcactta	agctcaggag	gtagaggctg
20401	cagtgtgcta	tgattgcacc	atgcacctat	gcaactcagc	ctgggcaaca	caacaaaaca
20461	ctgactctaa	aaaaacaacc	aacaaaaaaa	aatcacatgt	attcactggc	cctctctttg
20521	gggacctgct	acatagaatg	gttttttgtc	cccagttcac	tgacatcagg	tatggctatg
20581	tggcttgctt	tagaccatgg	actttgagtg	gaaatgacat	gtgccacttc	cacgaggaag
20641	ctttaaaagc	cgctcatggg	tctgccacct	tctctctctt	cgggtgtctg	agacggaag
20701	tccagcttg	agacttttcc	ttcagacagg	gctctRgaat	gaagatagca	tagaacagag
20761	tgggtcccatg	gaggacatgg	atatgagtga	gaaatcaaca	tgggtgtgtg	agcccctaag
20821	atttggggggc	tgctattact	gcagcgtaac	tggatcccag	ctgatatag	cagcctccct
20881	gtgggatacc	ctgctcaggt	atcctttccc	atccacctga	caactgacac	accataatga

20941	gctatgctga	tgtaggaag	tctccgcctt	tgtcctctt	cagagctgtt	caccctcagg
21001	tcctaaccag	tgagcctatt	tctttttttt	tttctttYtt	tttttYtttY	tgagaWggag
21061	tcttgctctg	tcaccaggct	ggagtgcagt	ggtgcgatct	cggatcaatg	caacctctgc
21121	cttctggatt	taagcaaata	ttgtgcttca	gcctcctgag	taggtctgga	actcctgacc
21181	tcaggccatc	cgccagcttt	ggccttctaa	agtgtctgga	ttacaggcat	gaaccaccgt
21241	gcccagccaa	gccgagtctt	cttgattctt	gctggcattt	ggcaactagt	agcagctgct
21301	cacaggaact	gtaaaaacat	ctgggtgggc	ccagaccttc	tagcatcaac	atgggtgccta
21361	gtaaatatca	atctcacatg	catcctgaga	tgcattaaaa	agaagctgtc	caggccgggc
21421	acgggggctc	acgcctgtaa	tcccagcact	ttgggaggca	gaggcggtg	gattgcttga
21481	gcccaggagt	ttgagaccag	tctaggaaac	atggcaaaat	cctagctcta	tttttaaaaa
21541	gggggggaaa	aagaaataaa	aaagctgggc	atgggtggtc	acacctgtaa	tcccacact
21601	ttgggtggct	gaggcagggtg	gatcacttga	gagaccagcc	tggtcaacac	catgaaacc
21661	catctctact	aaaaatacaa	aaattagcta	cacctcatgg	tgcaNcctg	tagtcccacc
21721	tactcgggag	gctgaggcag	gagaatcgct	tgaacctggg	aggtggaggt	tgcagtgagc
21781	ccagatcacg	ccactgcact	ctaacctggg	ctagagagtg	agactctgaa	aaaaaaaaaa
21841	aaaaaaaaaa	gagaaaagaa	cataatgttt	ggccaggcat	ggtgccttac	acctgtaatc
21901	ccagcagttt	gggaagccga	ggggcgggat	cacctgaggt	tagttcaaga	ccaacctaat
21961	caacatggtg	aaaccataat	ctactaaaaa	aaaaaaaaaa	attagccagg	cgtggtggtg
22021	gatgcctgaa	atcccagcta	cttgggaggc	tgaggcagaa	gaattgtttg	aacctgggag
22081	gcagagggtt	cagtgaaccg	agattgtgtc	actgcactcc	agcctgggcg	acaagagtaa
22141	aactccgtct	caaaaacaaa	acaaaacaaa	aaagaatcat	aatgggttagt	aagtgaataa
22201	tctgaattag	tttgtgtatg	tgtattgttg	catataatag	agacccaaat	taactgtggc
22261	ttaaataaga	tagaagttaa	tttctctctt	ctataaaagt	ccaagttagt	atgatggatc
22321	tttccatgaa	atcattagga	gccagatttt	ttgtatcatt	cattcattca	ttgattcatt
22381	actaccatta	atagagacaa	ttttctgcac	cattcaggct	ggagtgcagt	ggtgcaatca
22441	taattcactg	taacctcaaa	atcctgggct	ccagcgattc	tctgcctta	gccccacaa
22501	agtagcaggg	actacaagca	catgccacca	cgcctggcta	atttttcttt	ttcttttttg
22561	tagagggtgg	gtgttactat	gttgcccagg	ctggtctcaa	actcctggcc	tcaagtgatc
22621	ctcctgcttc	accctcccaa	agctctggga	tgacaggcat	gagccactct	gccccccag
22681	gtctttttta	tcttgttgc	gttccatccc	tagggcggtg	ccctcaccca	catgatccaa
22741	tatgattcac	caccacttcc	acagctgggc	ccttctgagg	ggtgatggtt	ggtcctttgc
22801	cctaaagagc	atgattcaga	agtacagatc	atttttgctc	taatcccat	agccaggatg
22861	tagtcatatg	gctacatccc	gatgaaagtg	ttgctgagaa	atagaatctc	tacctgagc
22921	agctttttgc	ccagataaaa	gttcagttac	tctgggagaa	gggtagaatg	gatactgggg
22981	gaccataagc	tggtgccacc	acacacatg	aatgttaacc	catcccaact	gtatcaattt
23041	ttccttctct	tcttctcttc	ctccctccct	ccctccctcc	ctccctcctt	ccttctcttc
23101	ttccttctct	ccttctcttc	tttgttctct	tctttcgaca	gtctccctct	atcccttagg
23161	ctggagtgcR	gtgttgccat	ctcggctcac	tgcaacctct	gcctcccagg	tccaagcaat
23221	tctcctgcct	cagcctcctg	agtagctggg	attacaggcg	tgctccacca	tgcccagcta
23281	atttttgtat	ttttagtaga	gacaggattt	ccccatgttg	gccaggctgg	tcttgaaactc
23341	ctgccctcag	gtgatccacc	cacctcagcc	tccaaaagtg	ctgggattat	aggcgtgagc
23401	cactgccttg	gcctcaaacg	gtatcaattt	tctgttactg	atttaaccaa	ttatcataca
23461	ctcagtgggt	taaaaccaca	cacatttact	ttcttacagt	tctggaagtc	agaagttcaa
23521	aatcagtttc	attgagccaa	tgtctggtgt	cagcagggct	gggtttttgtt	ggtggctctg
23581	gtggacaatg	tttcttggc	ttcttcagct	cttttttttt	tttttttttt	tgagacaggg
23641	tctcgtctg	ttacccaggc	tggggtgcag	tggtgcaatc	atagctcact	gcagcctcca
23701	tctcccaggc	tcaggcgatc	ctcccggtgt	agccttctca	gtagctggga	ccacaggctc
23761	acgccaccac	gcctgctaa	tttgttttat	ttttttaga	gatgaggtct	cactccattg
23821	cccagactgg	tctcaaagtc	ctggattcag	gagatcctcc	tgctcagcc	tcccaaaggt
23881	ctgggattac	aggtgtgagc	cgttgacccc	cacctctttt	cagtttagaa	aggctacctg
23941	tattccttgg	ctgggtgggtc	catcctccat	tggaagcac	atgaatccat	ctctgccttc
24001	atcatcactc	cactttctcc	tctgagactt	attcctcctg	tgtgcctctt	aggaggatgt
24061	tcatgattac	ataccgccct	cttgataaat	cctgaataat	ctctccatct	caggatcctt
24121	cacattttca	aaatcccttt	caccatataa	cgtgacattc	acagattcca	ggaataggac
24181	ctagacatat	ttaggggggt	tctctattca	gcctactgta	ccatgccatt	ccacacttaa
24241	gtgccttact	catttattca	taaaatatgt	attgagcaag	acctgtgtgc	caggcattgt
24301	gttaggtgct	agagaaatag	aggtgaaaa	acagacaagg	cctctgcttt	catggagttt
24361	atattctagt	gaagaggaca	agtaaatagc	taagctattc	tttttttttt	tttttttttt
24421	tttgagacgg	agtctccctc	tgtcgccag	gctggagtgc	agtggcgcaa	tctcggttca
24481	ctgaagccc	cacctcctgg	gttcacgcca	ttctcctgcc	tcagcctcct	gagtgcctgg
24541	gactacaggc	gcccgcacc	acgcccagct	aattttttgt	attttttagta	gagacgggt
24601	ttcaccgtgt	tagccaggat	ggtctcgatc	tcctgacctc	atgatccacc	cgctcggcc
24661	tcccaaagtg	ctgggattat	aggcgtgagc	caccatgccc	ggccaagagc	taagctattc
24721	taagctataa	cgtgtattat	caaaaacaatt	aaggccaggc	acagttgtctc	acacctgtaa

24781	tcacaacact	ttgggaggct	gaggcggtg	gatcatttga	ggtcaggagt	ttgagaccag
24841	cctggccaac	atgggtaaaac	cctgtctcta	ctaaaaatac	aaaaaaatta	tccagggtgtg
24901	gtggtgcatg	cctgcagtc	cggctactcg	ggaggctgag	gcacaagaat	aagaattgct
24961	tgagtgggga	ggtggagggt	gcagtgcagcc	aagatcatgc	cactgcacta	caggctagga
25021	gacagagWga	gaccctgtct	taaaaaaaaa	gcaattaggc	caagtgcagt	ggctcatgcc
25081	tgtaatccca	gcactttggg	aggccaagga	gggcagatca	cgagggtcaag	aaatcgagac
25141	cagcctggcc	aacatggtga	aaccctgtct	ctactaaaaa	tacgaaaatt	agctgggtgt
25201	ggtggcgctg	gcctgtagtc	ccagctactc	gggaggctga	ggcaggagaa	tgccctgaac
25261	ccgggagggtg	gaggttgcag	tgagccgaga	tcacgccact	gcactccagc	ctgacgacag
25321	agtgggaatc	catctaaaaa	aagaaagaaa	gaaattggct	ggagaatcgc	ttggaccag
25381	gggtggagggt	tgccatgagc	tgagattgtg	ccactgcact	ccagcctagg	caacaagagc
25441	aaaactccgt	ctcaaaaaaa	aaaaaaaaaa	tcccagcact	ttgggaggcc	aaggagggca
25501	gatcacgagg	tcaagaaatc	gagaccagcc	tggccaacat	ggtgaaaccc	tgtctctact
25561	aaaaatacaa	aaaatttagct	gggtgtggtg	gcgggtgcct	gtagtcccag	ctacttggga
25621	ggctgaggca	ggagaatggc	atgaacctgg	gaggcgagc	ttgcagttag	ccgagatcac
25681	accactgcac	tccagcctgg	gcaacagagc	aagactctgt	ctcaaaaaaa	aaaaaaaaaa
25741	gaaaagaaaa	gaaattaaac	agtgtgatgt	gacaaaaagt	gatagggggt	tggagacagc
25801	ttttctgttg	gatggttagg	aatggcttct	tagaaaagat	gactgacaca	tgggaggctg
25861	atgtggcaga	tcacgagggtc	aggagatcaa	gacctcctg	gctaacacgg	tgaacccccg
25921	tctctactaa	aaaatagaaa	aaattagccg	ggtgtggtgg	cgggcgcctg	cagtcccagc
25981	tactaaggag	gctgaggcag	gagaattggc	tgaacccggg	aggcagagct	tgcatggagc
26041	tgagatcacg	ccactgcact	ccagcctgga	cgacagagcg	agactccatc	tcaaaaaaaa
26101	aaaagaaaag	aaaagatggc	tgacacagag	ggcagagctg	agagccaaga	gggcagaaaa
26161	gagccataga	aaaccatttc	caggcctgga	agcctaagg	aatttcccag	ctggatttgc
26221	agttgctttg	gattggtgac	tcctttttac	ctttcattgt	taggggacct	gcaggttcct
26281	ttgcctcgtg	tgagctaca	gctccattac	accaagacaa	tagggatgca	gcagatgag
26341	ttactggtgc	agggcacctc	gtgcagagat	gggaagaggc	cctcaaatct	atctccccga
26401	gcaattcttg	gagagggttt	ctaaggggac	tgtggagggt	aggggattgt	ggagggttaag
26461	gttttgggca	actgggtcat	tgattgattg	ggggaaggat	gtagaagctg	cgtttttggg
26521	ggaattagct	ccttgtgggg	tccttcagg	cagctgagtc	agtagttcca	tgaggacctg
26581	aaggaatctc	ttttcttttc	ttcttcttct	tttttttttt	tttttttttt	gagatggagt
26641	ctctctctgt	cgccaggcta	gagggtcagg	gggtcgcagg	ctagagggtgc	agtggtcatga
26701	tcttggctca	ctgcaacctc	cacctcccgg	gttcaagcaa	ttctcctgcc	tcagcctccc
26761	aagtagctgg	gactagaggt	gcgtgccacc	acaccagct	aattttttgt	tttttagtag
26821	agacagggtt	tcaccatggt	ggccagggtg	gtctcgatct	cttgacttcg	tgatcggccc
26881	ccgccccacc	ctcggcctcc	caaagtgcgt	ggatcacagg	agtgagccac	ggtgcccagc
26941	cttaattttt	gtattttcag	tggagacggg	gtttcaccat	gttgatcagg	ctggagtga
27001	atggtgcaat	cttggtcac	tgcaacattc	gcctcctgga	tttgaatgat	tctcctgcct
27061	cagcctccca	agtaactggg	attacaggaa	tgcgtcacca	cgcccggcta	attttgtatt
27121	tttttagtag	agacgggggt	tcaccatggt	ggtcaggctg	tcttgaactt	ctgacctcaa
27181	gtgatccacc	tgtcttggcc	tcccagagtc	tgaagggaata	tctcaaagg	aacacttaat
27241	gttgtgtaat	gtccagggtg	tgatccatag	agcagttaaa	ggtaaaggta	actataattt
27301	tttttttttt	tttttagaca	gagtcctcct	ctctgtcacc	caggctggag	tgcatgttca
27361	cgatctcggc	gcactgcaac	ctccgcctcc	ctggttcaac	caattctcct	gcctcagcct
27421	ctcaagtgtg	tgctgccatg	ccaggctaat	tttttttttt	agacggagtc	ttgctctgtc
27481	accagggctg	gagtgcatg	gcacaatctc	ggctcactgc	aacctccg	tcctgggctc
27541	aaacaatgtg	ttttttcccc	tagtactttg	gtgtttgatt	atcttttttt	tttttttttt
27601	tttttttttt	agatagagtc	tcgctctgtc	acccaggctg	gagtgcatg	gtgcaatctt
27661	agctcactgc	aagctctgcc	tcctgggttc	atcccatct	cctgcctcag	cctcccaagt
27721	agctgggact	acaggcacc	accaccagc	ccggctaatt	ttttgtattt	ttagttaga
27781	cggggtttca	ccatgttagc	caggatggtc	tcgatctcct	gacctcatga	tctgccgccc
27841	tcagcctccc	aaagtgcctg	gattacaggc	Rtgagccact	gcacccagcc	tgggtgttga
27901	ttatctatta	tgtcaaacag	gctgggttta	gtggctcacg	cctgtaatcc	cagcactttg
27961	ggagactgag	gtgggaggat	cacttgagcc	caggagctga	agaccagcgt	agcaatgtag
28021	caactccctg	cctctacaaa	aagttaaaaa	atttagctgg	gtgcaccagY	agaccagctc
28081	ctcaggaggc	tgaggaggga	ggatcactcg	agcccaggag	ttcaaggctg	cagtgacgtg
28141	tgatcatgcc	actgtactcc	agcccaggca	atggagcgag	accctgtctc	aaaataaata
28201	aaacatgaag	aatgtcgaac	acattatctg	gtttttgttt	ttgttttctt	tttttgagat
28261	gttgtctcgc	tctgtcacc	tggctggagt	gtagtgggtg	gatctcggct	cactgcaacc
28321	tctgcctccc	gggttcaagc	gattctccc	cctcagcctc	ccgagttagt	gggacacag
28381	tgactgtccc	ccatgcctgg	ctaatttttt	ttattttttt	tttttcagta	gggacagggt
28441	ttcgccatgt	tggccaggct	gttctgaaac	tcctgacctc	agatgatcca	cccacctcgg
28501	cctcccaaag	tgctgggatt	acagggtgga	gccatcgtgc	ccggcctggt	ttaaaaaacc
28561	atattggccc	aactcgggtg	ctcatgcctg	taatcccagc	actttgggaa	gccaagcag

28621	gaggattggt	tgagcttagg	agtttgagac	cattctgggc	aacatggtga	aacctgtct
28681	ctgcacaaaa	atagaaaaat	ttgccacctg	tgctggtgtg	tgccgttagt	cccagctact
28741	ctcaaggctg	agggaggagg	attgcttgta	gagcctggga	agtcggagct	gcagttagcc
28801	atgatcacac	caccacactc	tagcctgaca	gaatgagacc	ttatcccaaa	agaaaaata
28861	aatgatattg	tattatatgt	gaactttgaa	ttatattgtg	ttgtatctga	agtttgaatt
28921	ttcacgttat	gtttaaaaat	cttggtggg	cgtggtggg	cacgcctgta	atcccagcac
28981	tttcggaggc	caaggcgggt	ggatcacctg	aggtcaggag	ttcgagacaa	gcctggccaa
29041	catggtgaaa	ccccgtctct	actaaaaata	caaaacttag	ccgggcatag	tgacatgcac
29101	ctgtagttcc	agctactcgg	gaggctgagg	caggagaatc	gcttgaaccc	aggaggcaga
29161	ggttgacagt	agctgagatc	gtggcattgt	actccagtct	gggcaacaag	agtgaactc
29221	catctaaaaa	ataaaaaaga	aaaagaaaaa	ataatacaag	aaattagccg	ggcgtggtga
29281	caggcacttg	tagtccctcc	cagctactca	ggaggctgac	gcaagagaat	tgcttgaact
29341	tgggaggtgg	aggttgacgt	gagctgagat	cgtgccattg	caacttagcc	tgggaacaa
29401	gagcaaaact	cagtctcaaa	aataaatagc	ttgaaccggg	gaggcagagg	ttgcagttag
29461	ctgagattgc	accacttcat	tccagcctgg	gtgatagagc	aagactctat	ctctaaataa
29521	ataaataaat	aatcctttag	gatggcaatg	aatttaagga	ctaaactagg	gagaactcag
29581	tttttttttt	aaaatggagt	cttggtctgt	cgcccagact	aggggtgagt	gggcgccatc
29641	tcggctcact	gcaacctcca	ccttccagggt	tcaagggtatt	cttgtccctc	agcctcccaa
29701	gtagctggga	ttacaggcac	ccgccaccat	gcctggctaa	tttttgtatt	tttagtagag
29761	atgggggttc	accatggttg	ctatggttgg	ccaggctggt	cttgaaactcc	tgacctgagg
29821	tgactctgct	gcctcggcct	cccaaagtgt	tgggattaca	ggcatgagcc	actgtaccca
29881	gcccatttca	cattatttat	ttattttatt	atatttttat	tttttgaggt	ggagtctcac
29941	tctgtcgccc	aggctggagt	gcagtggcac	aatctcggct	caactgaaacc	tccgcctccc
30001	gggttcaagc	cgattctcct	gcctcagtct	cccagtagtc	tgagattaga	ggcaaccacc
30061	actatacccg	actaattttt	gtatttttca	gtagagatag	ggcttcacca	tggtggccag
30121	gctggtctcg	aactcctgac	gtcagttgat	cctccacact	cagcctccca	aaagtgtgga
30181	attaaagctg	taagccagcg	ggcctggtgg	acatctttta	ataatcagtc	tttccattca
30241	ggtatatggt	atatgtctcc	atttacttag	gtcttatttc	atatccttca	ggttggagct
30301	atcattttct	ttcatacagg	ttttgcacat	ttcttgtgag	gtttattcct	tcattggtcca
30361	tggattttgt	tgtgaattgg	gaatcctttt	tccaccaagt	atatttttcta	atttgttact
30421	ttagtataca	ggaaagataa	ctaattttta	tctgcagttt	attatctatg	aaaggaataa
30481	agtagaacta	ctcagtaaaa	ggtttccata	atcaaataag	tatgggctaa	acaaagctaa
30541	acagatgtgt	tcactgctgg	acttatcaat	gcttgtgata	atattttttt	ttttttttga
30601	ggcagagttt	tgctctgtag	cccaggctgg	agtgcagtg	cgggatcttg	gctcactgca
30661	acctccacct	cccgggttca	agtgaattct	ctgcctcagc	ctcccagata	gctgggacta
30721	tggcatgcac	caccacatct	ggctaatttt	tgtaatttta	atagagacag	ggtttcacca
30781	tggtgactag	gctggtctca	gaactgttga	cctcaggtga	tctgcctgcc	tcagcctccc
30841	aaagtgtctg	gattacaggt	gtgagccacc	accaccaggc	aattgaagac	gtatatttcta
30901	tgaagaaatg	ggttagatttt	aatgaacaat	accccttttg	tgggcagatt	cctaagtcct
30961	aggccctcac	aacaaagggg	cagtgggcct	ggagatgcca	gcttcagctg	ccagagggac
31021	tgctcctcca	gggccacccc	agccacattt	tgatcaccaa	gttttgatca	ccaagaatcc
31081	caagaagggc	acagggaatt	tcctttctta	cctgcccatt	aaaccttttg	tcactagaca
31141	tcctgaaaca	tactttggga	aactgcattc	aaagaccctt	ctagtttcaa	atctgtggat
31201	ccagggggtc	ccactgaacc	ttacctgatg	cccaaactcc	caccatttca	ctcccaacca
31261	gaacacagaa	gatgacctgg	tgccaaaatg	aaagctttaa	tgagtgttac	tcctagacag
31321	tcacgtctca	gcttctgcca	gcctccactg	tcccagctct	cttagctggc	cgacagggga
31381	gctagtgtgt	gaggggtagg	gatctggagt	ctaaagagca	gagccaggca	aaaggaggta
31441	caggaagccc	ccgatggggg	ctgggctccc	ggagtgtggt	gctggggggg	catgggcttc
31501	aggccggccc	ctcttcaggc	attcctagca	aagccaccag	gggctccagg	ggtgtggggg
31561	tccccatggg	cacagggtgg	gtgcgttcat	gcttgcgcaa	gtcgttgcca	ctcaagaagg
31621	ccttgggaca	atgggggcag	gtgtaggggc	gcactgagct	gtgagtgcgg	ctgtgtttgc
31681	gcagcccagc	ccggtcagag	aagctcttgc	cgactRggt	gcaggggaag	ggccggagct
31741	ccgggtgtga	gcgctcgtgc	cgacgcagca	gcgtcattgt	ggagaaggtc	tccttgcact
31801	ctcggcacac	aaactggggg	ggcttctcgt	cagcctcctc	accccccgcc	tcgctgccc
31861	cttccaaggg	accaggagcc	tcccgagcac	cagcatcttg	gcattccaca	tgctccaccg
31921	tcactcccac	cactgcccac	tgtgtggcca	tcacaccacc	tgactccggg	ggcagcccta
31981	gcagccctgc	tggagggtcc	cccagccccg	cccctgctRc	cggggcggct	gaactctcac
32041	ctgccacgcc	cacaggcagc	gccaaaccca	ccaccagctc	ctgtgcaggg	ggcacacccg
32101	cggcctcact	gcttcgatgg	gtccgctcgt	gcttcctcag	gctcgacgac	accacaaagg
32161	atttcccaca	tgcgttacag	tggaaagggc	gctcccccca	gtgcacccgg	ctatgcttcg
32221	tgaggctggc	acgctcggcg	aaggctcgcc	cgcactcctc	acagcggaag	ggcgcgtggc
32281	cagagtgcac	cagcgctgc	cgcttgaggt	cccaggaYgc	cacgaacgtc	ttgtcacatt
32341	gcaggcactt	gaacggccgg	tcgctgtgt	gcacacggcg	gtgcatggcc	aggtccggcg
32401	gctgccggaa	gtccttgccg	cacttctcgc	agtgttatgg	cttcaccctt	tcattgggcg

32461	gctgggtggcg	acggaagctc	gagggggtcgg	agaacatgcg	gccgcagcgc	gggcacagga
32521	agggcttctc	ccccgagtcg	gtgcgctcgt	ggctctggta	ggaactgagc	tgctggaagc
32581	ccttgccgca	ggccgggagc	cggtagggct	tctgtgccgc	gtggatgcgc	tggtggcacg
32641	tgagcgagga	tgagcgggag	aagctcttcc	cgcactcggg	gcagaggaaag	gggcgctcgc
32701	cgggtgtggga	cctgcggggg	tgtggaggac	ttggcatgaa	ggcgacagac	ccataacgtg
32761	acccactgc	ctgtctgggc	tgtactttag	gggctcccca	aacgttcgtg	ggggcctagg
32821	cttaatcccc	taagagccac	atggctgcac	cccagaggaa	gaagccttca	ggctggctgg
32881	gtgtctctat	tccaaagacc	tgtctctgca	cattaaagac	caagatatgg	gccggggcgc
32941	gtggctcacg	cctgtaattc	cagcactttg	ggaggccgag	gtcaggagat	cgagaccatc
33001	ccggctaaca	cggtgaaacc	ccgtctctac	taaaactata	aaaaattagc	caggcgtggt
33061	gggtgggtgca	tgtagtcccr	gctactcggg	aggctgaggc	Rggagaatgg	cgtaaccgcg
33121	ggagggcgag	cttgcaagtga	gccgagatct	tgccactgca	ctccagcctg	ggcggcagag
33181	cgagactccg	tctcaaaaaa	aaaaaaaaaa	aaagaaaaag	aaaaaaaaaa	aagaccaaga
33241	catggccagg	cgcggtggct	cacgcctgta	atcccaagcac	tttgggaggc	cgaggcgggc
33301	ggatcacctg	aggtcaggag	ttcaagacca	gtgtgaccaa	aacggagaaa	ccccgtctct
33361	actaaaaata	caaaattagc	cggcatgggt	ggcgcatgcc	tgtaatccca	gctactcggg
33421	aggctgaggc	aggagaatcg	cttgaacccg	ggaggcggag	ggtgagggtg	gccgagatcg
33481	cgccattgca	ctctagcctg	ggcaacaaga	tcaaaactcc	gtctcaaaaa	acaaacaaac
33541	aaaaacaaaa	caaaacaaaa	cagaaaacca	agatacgtgt	cctccgcctt	tttttctctg
33601	ttccccaggc	tggaaatgcg	tggcctgacc	atagctcact	gcagcctcga	cctccagggc
33661	tcaggccatc	ctccacctt	atctcccaa	gtaccgggga	ctagaagtgt	acatccccc
33721	gctcgggtaa	tttttttatt	tttatagaga	cgaggcttgc	tgtgttgccc	aggctggtct
33781	tgaactcttg	ggctcaagca	atcctcctgc	ctcagcctcc	caaagtgtctg	gaattatagg
33841	cgtagagctat	tgtgcccagc	ctagaaacat	gtcattaatg	tagaggctga	gaaaaagaaa
33901	aaaaaaaaatg	acctagacaa	accaggcccc	actcacacct	cctggtctcc	acaaaagacc
33961	ctcagaactg	cccaactcca	aaccccgccc	cctttccagc	tggcctacaa	cggaggccaa
34021	tctgacccaa	tcccatcttc	agagatcaac	ctcaagggtg	ttgccacctc	tgcccaatca
34081	ggggcaccaa	tttctcccac	atgcctagcc	cctccccttg	gatctgccat	gccaccttc
34141	ccattggctc	actttaccct	gagactcaaa	cccaggcccc	attggctgca	gcaacgctgt
34201	cgccctgccc	cgggaaggcg	cctgcccccg	aaggcRccct	caccgctcat	ggttgccgag
34261	gtccttgagc	tccgcatagg	ccttRccgca	acgctcacag	ctgtagggcc	cgaggccagc
34321	gtgagtacgc	cgggtgcttg	ggaacactga	agggtcagca	aagctcttgc	cgcagtcggc
34381	gcaggcgtaa	ggccgctcgc	ctgtgtggcc	acgctggtgg	atcttgagct	tggagagcgc
34441	gccataggcc	ttcgggcagt	gcgcacagcg	gaagggcagt	tcgccagcgt	gcgaggccag
34501	gtgcacgcgc	aggcacacgg	gctgcatgaa	gcggcggccg	cactcggggc	acggaaaagg
34561	cttctccccc	gtgtggctgc	gccggtggct	gcgcagctcg	ggtgccgtct	tgtaggcctt
34621	ggggcatagc	ggacacgcac	agggcctagg	cttggcccg	gagcctgaca	ccttctcccc
34681	actggcttcc	tctgccttag	cttctgtctc	tggctttggc	ttcacctcgg	ccacctcttc
34741	agagcagctc	gccggcccat	gtgtggcagc	gtggcgcgct	gccttggg	cgtttgaaa
34801	tgtcttggtg	caggacaSgc	acttgtagcg	gcggcccag	cgctttagc	cgggggcttg
34861	ggaccggg	tctgcagcct	ccacttccat	ggccttggtg	aacggggtt	ctctgcaaga
34921	gaagcaaa	tagaccaa	ccacatacct	tcgccactcc	tgaagcctc	agagagaacc
34981	ctatctcatc	tgcatttctc	acctcggaac	ccacacatcc	ttcctgcca	gcattcctgg
35041	ctctgacatc	ctgcgttcgt	ttcctccctg	atctgtctcat	tgaagaaa	agttggacca
35101	agtgtccgca	gagccactaa	gaaaggaggc	tgagggtcac	aaaagattca	cctacacgtc
35161	cccccccScc	cccaacgggc	ttttccaaac	actgtggcat	tcccagaggc	ccaggttcca
35221	tctgtctcac	catcttccct	cttcagctca	gtgtccaaga	actatgccag	gataaagagt
35281	gtaccagac	gtgggcctgg	cctgaaggtc	tccaagcgcc	cagaaaagac	agacctggga
35341	ccagaaaagg	gctgagccaa	tgggctaaat	ctggtagctg	gcactgtctg	gaagtgcag
35401	gtccccagca	tttgtgtttt	ctttcctcct	ctggatggtt	agtcctcaga	gacagcaact
35461	gttcacacag	aattctggcc	ttgcacagct	gtacgggctt	ccgccccaga	ctggaatctg
35521	tccactctct	gctctggaat	cttgttggcc	tgttcccaca	actctggtaa	tggagaatca
35581	ctcaaggcag	cctgagccat	tgctagcagc	tggaaagcctc	tttctgagtc	ataactgatg
35641	tatctgatct	aacatggcct	cctgggatac	cagctctagc	tgagatccct	actttctggt
35701	ccagaaccga	gacgccttcc	accagctgc	tcctggggat	catggttggg	aggaaacagg
35761	attaatggct	gtattagtct	taacaccagc	tcactctccc	tggggagtga	aggggaagag
35821	attatggcag	atccacttaa	ggagtgtca	gcagctgctg	ctgggggaag	gggtctgagg
35881	agtgggggct	gcagggagcc	agggttgccc	agaggctagg	gggcctacgt	tctacttgca
35941	gccctgtgga	ttactatgag	acctcagtga	aatgagtgtt	gtttataaga	ctatttccgc
36001	ccggctgggc	gtggtggctc	acgcctgtaa	tccagcact	ttgggaggcc	gagggggg
36061	gatcacgag	tcaggagacc	gagacatcc	tggctaacac	agtgaacccc	cgctctact
36121	gaaaatacaa	aaaaattagc	cgggcgtggt	ggcgggcgc	tgtagtccca	gctactcggg
36181	aggctgaggc	aggagaatgg	tgtgaacctg	ggaggcggag	cttgcaatga	gccaagatcg
36241	tgccaccgca	ctccagcctg	ggcgacagag	caagactccg	tctcaaaaaa	aaaaaaataa

36301	aaagactatt	tcctcatctg	gacacgttag	gggttggtgg	cttctaagtt	atgacactgg
36361	ggttcaggag	gaggaaactg	agctgagcca	cgagggtgct	aggggaaata	cagagactca
36421	ggggccctta	tgccaggaaa	ggggcgggaga	agcagcttac	ctgggtgacc	caggaggagac
36481	acaggagccc	ctgacgcgtg	gcaagggcca	catccttaaa	gtcagcaccc	ttctgcctga
36541	aatcccggcc	tcaggcctgc	acgctgcccc	ccctcccca	ccccatgcag	ccggtcttct
36601	cccaaaaagta	atgaatgatg	tttccctgttc	cctgctcaag	aactttccat	ggcttccacc
36661	acaccatgca	aatccccagg	tccttccctgc	ttgttccaaa	tctgacaagt	ccttcaactgg
36721	agccccactt	ccaccaggag	gatcctcaag	catctcccct	ttgggttttg	ccccaatccc
36781	acaaaactcc	atcttattct	cacttggtta	tttactttt	cccactaaat	cctagggtct
36841	agccagggtg	ggtggcatgc	gccttaagtc	ccagctacgt	gggaggctga	ggcgggagga
36901	tcctgtgaac	ccagggggttc	gagcgtgcag	ggaactatga	tcgtaccact	gcgctccagc
36961	ctgggcaacc	tggtgagacc	ccatctctac	taaaaataaa	aattagcaga	tgtggtggca
37021	tgtgcctgta	gtactgccta	cttgagaggc	tggttggtgga	ggactgcttg	agcccaggag
37081	ttcaaggctg	taccactgca	ctccagcctg	ggagacaggg	caagacactg	tctcaatcaa
37141	tcaatcaatc	aataatcaat	cctggggcct	gaagataagt	taaagggact	gaattctaac
37201	ctttctgatg	acttgaattc	ttcctacagt	ttccaaggga	tccctcccta	tttctggatg
37261	aggtaactac	tacctcttcc	agacggttcc	tggagagtct	gcctgataat	gttccctcct
37321	aataaaaaatg	atagcttggg	ccaggttcag	tggctcacac	tagcactttg	ggagaccaag
37381	gcgggtggat	cgcttgagcc	caggagtcca	agacaaggcg	ggtggatccc	ttgagcccag
37441	gagttcaaga	ccagcctggg	cgatatagca	aaatcccatt	tctacaaaaa	atacaaaaaa
37501	tagccaggca	tggtggagca	tgctgttact	cccagctact	ccagaaggct	gaggtaggag
37561	aaatgcttga	gcctaaggag	actgaggttg	cagtaagcca	agatggtgcc	actgcactcc
37621	agcctggcaa	cagagtgaga	ccccatctca	aaaataaata	aataaataaa	tgataaaaaa
37681	gatagtccac	atttactgag	cactcgccaa	ataccaggca	gtatccctaaa	ctccttatgt
37741	gtattagctc	agttaccctt	catggcaacc	ccatgaggaa	agttctatta	ttccattttc
37801	acagataagg	gaaccaagg	ccagatggag	ggttcagtat	tttgttaagt	gcccagctcc
37861	tgaagccaaa	ctgtctggct	tcagattttg	cctccatcac	ttcccagctg	atgtgaccgt
37921	gtgtaatgta	ctgcatgtct	tagaacctca	gttttcta	ctgaaaaatg	gagataatga
37981	Yagtacttac	ctgacagagt	tgggtgagga	atgaatgagt	caaaaaataa	tactgtcctc
38041	aattatcaaa	gcgtcttctt	agagccagac	acattgctgg	gtgtgctggt	ttattttaatt
38101	taattgtctta	taScttYctg	agatagggat	tcttgtccct	actttacaga	agataaaatc
38161	gaggttccaa	gagtttaagt	acttgcccat	ggtcccaaca	tgggtaacct	aagcagctgg
38221	gacttcagtc	caggtgttta	atttgcttta	agttgctggg	gtcttgctca	gtggtctggg
38281	gcctcttacg	cttgtctgct	gcctccgcca	gccccacagt	gaccagaacc	ctgagctcag
38341	gtcatacctg	tgtcttctct	catccttgca	gaactgccct	gagaccctgg	ccggcacctc
38401	ttatgtctct	gcttccctct	cagagggtct	gacccaagtgg	ttctgagctc	tggcccttct
38461	actgcctctt	gccagctctg	Kgtctcagcc	ttcctatctg	tgagttagac	accaggtagc
38521	tggaggggaa	atccctcctc	ccatggcact	tcccagggga	aaaggtaggg	gagtgccagg
38581	ttgggtctcag	catgcgccca	gctacacaaa	gaggScagg	aggctagggt	tctgtctaac
38641	atcccaccat	taaaaaaaaa	aaaaaaaaaa	atatatatat	atatatatat	atatataatt
38701	tttgagatgg	tactcttactc	tgttgcccaag	gctggagtgc	agtggcgcg	tcttggttca
38761	ctgcaacctc	cgctccttga	gtagctggga	ctacaggcac	ctgccaccac	acctggctag
38821	ttttttgtat	tttttagtaga	gacgggtttt	cactgtgtta	gccaggatgg	tctcaatctc
38881	ctgacctcgt	gatccgaccg	cctcggcctc	ccaaagtgtc	gggattacag	gcgagagcca
38941	ccaMgcccag	actttttttt	ttttttttat	taaagagctt	gaggtaggcc	tcagggaattc
39001	gtatttttaaa	tacactctgg	atcattccag	ccaatacttt	tgtttgtttg	gttttgaaac
39061	aagatctcac	tctgtcgctc	aggctggagt	gcagtgggtc	agtcatggct	tactgcagcc
39121	ttgacttccc	agactcaagc	aatcctccca	cttcagcctc	ccaagtagct	gggactacag
39181	gcatgcacca	tcatgcctgg	ctaattttta	ttttattttt	agtagagatg	aagtcttgct
39241	atgttgccctg	ggctgggtcta	gaactcctgg	gctcaagtga	tcctcccMcc	gcagcctccc
39301	aaagtgtctg	gattacaggc	gtgagccacc	gtgcctgccc	aaccaatact	taagaaccaa
39361	acacacatcc	ttaggtctcc	acgagctctc	aggagaggag	catttttaagt	gttcaactaca
39421	cctctttttt	agatattgag	attaagggtcc	ccacaaagga	aaaactgtac	acaaggacac
39481	acagctggtc	aaggagccag	actcgaaccc	aagtctccat	tctctcccc	aggttgaatc
39541	atgagacttc	ccactgctcc	caggaaaaag	accaatatct	tttccatggc	cagcatagcc
39601	ccaaccatc	taaatcctgc	ctactgggc	agatcacttg	aggccaggag	tttgagacca
39661	gcctggccaa	catggtgaaa	ccccatctct	actaaaaata	cacacacaca	cacacacaca
39721	cacacacaca	cacacacctg	cctacctcac	ctccactcc	tctccctggc	ccactgggct
39781	ctaccacacag	aggcctcctt	tcttctcctc	aaagagctaa	attccttccc	acctcagggc
39841	agtggcacta	gcagttccct	ctgtctgagc	cactcttctc	ccacgatctt	tgtgtagctg
39901	tctttttttg	tggtattttg	atctcagctc	ccagtcaccc	cctcaaaaag	agctttcttg
39961	accacctttc	cttttcttcc	ccccctttta	tattccaaat	tttttccctt	tttaaccaac
40021	caaggagcac	tgaatgacta	cctttctcaa	tgctatcttt	acccctgata	atcattctct
40081	atctactctt	tattattatt	attatttttt	gacggaatct	catgattatc	tatcaagcac

```

40141  ttctcctgcc  tcagcctcct  aagtagctgg  gactacaggt  gcccgccacc  acgcccggct
40201  attttttttt  tttttttttt  tttgagacgg  agtctcactc  cgtcaccag  gctggagtgc
40261  agtggcacia  tcctggctca  ctgcaagctc  cgctcccgg  gttcatgcca  ttctcctgcc
40321  ttagcctcct  gagtagctgg  gactacaggt  gcccgccacc  acgcccggct  aattttttgt
40381  attttttagta  gagacaggg  ctccagtgt  tagccaggat  ggtctcaatc  tcctgacctc
40441  gtgatccacc  tgcctcggcc  tcccagagtg  ctgggattac  aggcgtgagc  cactgcaccc
40501  ggcccaatcc  cggctaattt  ttgtattttt  agtagagatg  gggtttcacc  atgttgacca
40561  ggatggcttc  gatctattga  cctcgtgatc  cgcccgctc  ggctcccag  agtgctggga
40621  ttacaggtgt  gagccaccgc  gcccgccct  ttttttgaga  cggagtctta  ctctgtcccc
40681  caggttgag  tgcaatggca  caatatctgc  tctactgaac  ctccgctcc  cgggttcaag
40741  cggttctcct  gcctcagctt  cccagtagtg  tggaattata  ggcccccgc  actacatctg
40801  gctcatTTTT  gtatttttag  tagagagagg  atttcacat  gttggccagg  ctgggtctga
40861  actcctgacc  tcaagtgatc  caccacctt  ggcttcccaa  agtgcttgga  ttacagggat
40921  gagccaccgc  acccagccct  ctttactttt  taaaaaatgt  ttttattttt  atttatatat
40981  ttatttttga  gacagagttt  cactcttgtt  gccaggctg  gagtgcagt  gcaccatctc
41041  tgctcactcc  aacctccgc  ccccgggtc  acgagatttt  cctgcctcag  cctcccaggt
41101  agctggaatt  acaggcatcc  accaccagc  ctggttaatt  tttgtattt  ttagtagaga
41161  tggagtttca  ccatgttggc  caggctggc  tcgaactcct  gacctcagat  gatccactgc
41221  ctccggttcc  cagagtgtg  ggattacagg  catgagccac  cgtgcctggc  ttatttttat
41281  ttatttttat  atttatttgt  attattattt  gagacagagt  ctccctctgt  tgcccaggct
41341  ggagtgcatt  ggtgtgatgt  catctcactg  cgacctctgc  ctccYgggtt  caagcaattc
41401  ttctgctcca  gcctcccaag  tatttgggat  tacagggtgc  tgccaccaca  gccagctaatt
41461  tttttgtatt  ttttagtag  atggccatgt  tggctaggct  ggtctggaac  tcctgacctc
41521  aggtgatcca  cccaccttgg  cctcccaaag  tgctgggatt  acaggcttga  gccaccatgc
41581  ccggcctatt  tatttcat  ttatttattt  attttctttg  agagaaagtc  tttgttgccc
41641  aggttgaggt  gcagtggctg  catctcagct  cactgcagcc  tccacctccc  ggattcaggt
41701  gatttctccag  cctcagcctc  ccgagttag  gggactacag  gcgaaagcca  tcacacctgg
41761  caaacttttg  tatttttagt  agagacagg  tttcaccaca  ttggccacgc  tggctctgaa
41821  ctctgacct  caagtgatcc  gaccgcctca  gcctcccaaa  gtgctgggat  tacaggcgtg
41881  agccaccgtg  cctggccttt  atttttattt  agagattggg  tctactctg  tcaccttgga
41941  gtgcagtggc  tcaatcatag  ctcaactcag  cctcaaaact  ctgactcaa  ctgacctcc
42001  tgagttagta  ggactatagg  caccaccac  cacacctgc  taatttatta  aaaattcttt
42061  gtatatagag  atggaggtct  cactacgctg  ccgacagtgg  tctcaagaac  tcctggcttc
42121  agatgtcct  cactWtggc  ttcctaaagt  gctgtgatta  caggcatgag  ccacagtgc
42181  cagcMcccc  cctctctaat  ttcctttatg  gtgtcatctg  gacaatactc  ctgtcaagct
42241  taccatgggc  aaggtatcat  tctaagcatt  tctgacataa  tactcaacta  ctcaagccaa
42301  ctgcacagct  gcctagcagt  tcattatgag  tgaatgtttg  tgttccctgt  ccRttcatat
42361  attgaaattc  taactctcat  tgtgactgta  tttggagaca  aggcctttat  ggaagtaatt
42421  aaggctgaat  gacRtcataa  gggataggcc  ctggtccagt  aggattagcg  ttcttatgag
42481  aagtgcacc  acaagaaac  aagcttacta  gtcacggctc  cagccagtgt  tcaaatccca
42541  aacacctgct  ctctgagccc  tgactattgc  tttgctagca  ttacacatct  tatggtttg
42601  ctgttaattc  ttcacacca  gcaccagagt  ccaggctggc  aaagggttag  gaaaccgatc
42661  atctgcctcc  tctacacca  gaacctgtg  tggtagccca  aaacaaatgg  aaagaacca
42721  cctcagatga  aatttgaacc  caggtctgta  gcctctgcct  cctccacagt  aggagtgtg
42781  gagaatgtcc  accaataact  gttatttaac  taaattcctc  cactttcca  actccacaag
42841  gctcaactat  tgcccaata  tcccacagt  ggctccctgt  cgtggcgatg  aagccttgc
42901  ttgcctcact  actggcttag  gaaggggatg  aagttctgct  gtctcactag  gcaagcaaag
42961  attcaatttc  caaaaatcct  aggtaggac  cctggggcag  ggatgaggag  aaaaaggagg
43021  cacctcaatc  tcccatctc  taaacaagca  gtcaccacac  aggcctcacg  gccaggagt
43081  acgtagttag  cggatgaacc  ctgcagaagt  gagcgtaat  catatcctag  gcctcctacc
43141  gggccagttc  acaacctgat  ccaacctac  ttttcagcc  tcagctccag  cgacagatc
43201  ctccctccgg  cagcgaccga  acctccagcc  acaattctca  atcctccacc  tctagatgca
43261  gtttccctgg  ctccggggag  ccttgctgc  cttctccaac  gggaaagccc  ctttacatcc
43321  ttcaagaacc  ccttccctta  gtgcctcag  gaatacttta  tggcagtaga  gaggttaagg
43381  ccccgccgg  ctctggacct  cagtttccct  atcagtaaaa  tggggtccct  taagtttgtg
43441  ggagtttgaa  ggccgggtgg  gcctacgaag  cgcccgagg  agccgagagt  tgcaagacc
43501  cggagctcct  cgctcctcgc  aaccgtctgt  acgcccgcgc  cccgccagcc  aggcagcccc
43561  tggagggcag  gaccccggtt  ataagcctca  gaaaacgtgg  cttcgaggga  cgtggcaagg
43621  aggactgaac  gaaggatgag  gagatgaaca  aatgaatgga  cggaaggccg  aatgagggac
43681  aaaggcttgc  aagatggcgt  tctctaggac  cgccggagtg  gtggccgggc  ctgacaggg
43741  agggggccg  ccggcgctg  ggtcttgca  gcgcccggca  cgccagccgg  acaaaaacc
43801  aaaacatggc  gggctctagg  acccccgga  ccacaccgcg  ccggccagc  agccgcgcgg
43861  gccgagcctg  ggtgtccgca  gccagaacc  gcggagacag  ccggcggtt  ctaggacctg
43921  ctggggccgc  aatgcgccc  gggcgctca  cagcccgccc  cgcccgct  gctcgctg

```

```

43981  ggccaacccg gcccgcctc ggcgcccga gggaaactga ggcagctcg gatcgtggcc
44041  gcgtgggagc tgcccggggc cctcgcggt tcccgcgac gtttcctacc tgatgagact
44101  tgtgtgact ccgtggcgtc ggcgtcggt cctcgaccg acggagcccg gaccctgcca
44161  aacaggggccc ggcgctagga cccagcggg ggcggggag tgggaccggt ggcgcggcga
44221  gcggaagtga gggatcttcc tcagctagga aggaaggaa agttcccggg gaacctccag
44281  cctatggcgg cagagaagca tcttgaaga ggtctctgtg tgtgtgagg caaggggacg
44341  ccaggcaggc tgacgggtata cgcgcgctt gtgttagtct gggggcaggc caccggcaat
44401  gtcttcaaga accagaaggt gggaggacaa aaaggccatg ctcaagctct gcaaaaccta
44461  agtccaatta tattccacaa catttccttt aatagcaagg agtgggtttc acgaagtttt
44521  ctccatctgc cttgggaaaa gtccctccaga gagccccacg aaaagtgtca ccaaagtacY
44581  gtggccaggg ctgtagaatc ttttctccc tttctgcgc tattcaatat gcaccggaac
44641  gatttctagc gctcttgccg gaagtgtggc tagcWtagcg tgcaaagacg ccgctgtgtg
44701  acctacgtgg gcggagtcag ccgtgcagac ctagaactta gcgcgggaag tgtgtctgcg
44761  taaaactgcg gctcggaggc gggacaggca gtgctcccga agcggaaagt tcgcggggaa
44821  gcttttgac ctaaggacac ttctgtttc ctagtaacaa taggaagtgt ccgtagagct
44881  gggagtagac tcctggctca tccagctgc gccctcttt ctctacctcc ttctcttcc
44941  ccccttctcc tttccctctt tcccttccct cccacctccc gggaacctcg gctgagtgtg
45001  cgtgtgtggg agcgcgagag ccccccga cccacctctt ggggcgcgc gctgcagtta
45061  gggtaagggt cccagtgcgc aggcgcgctc ttgttcgcg tcccactcg acgcgccgc
45121  ggcgggaagg agagggggcc gccggtgcga ggtaggcgct ctgagaggaa attggcagac
45181  gagatgagtg aggtcagagg acttctagg tggttgacta ggggagggga cttttagtca
45241  aaggggcttt gtgaggtgag aattccgggg gcaggttcag tgaagggtt tacggctcga
45301  gacagttgta ctgggtccc atgggaagaa gagtcagagt tggagaagga caagacgatg
45361  tagggggcat gagagagcaa gggctttggg atgaagaccg cctcagtcag ggggtttgct
45421  gttacgtgaa ctagaataac aggcattgcc gtgtgttctg aggattgagt aacacaatga
45481  atatagttag cacagtgcct ggcacatggt aatataatac tctcatgag ttgctgcat
45541  catcagttat aagagaggag gcagaagaaa aaaatgagaa agactggtga gggtaagatc
45601  aggaggtcat ggagagaaaa ggtaggcaat gaagtacat tacaacctgg tattgatgtt
45661  attcccaaga tggaagatag tttgagttca agggaagtag ataaaagaag tcgctaagag
45721  agtcttttgg ggttttttKt ttttttgaga tggagtctcg ctctgtcacc caggctggag
45781  tgctgtggtg ggtatcagc tcactcaac ctccacctcc caggttcaag cagttcttca
45841  gcctcagcct cctgagtagc tagaattaca ggtgccacc accacgcccg gctaattttt
45901  gtatttttag tagagacagg gtttcaccat gttgtccagg ctggtcttga actcctgaac
45961  tcagggtgat caccagcctc ggcctcccaa agtgctggga ttacaggcgt gagccaccac
46021  acctggccca gttaagagag tcttaactct ctttaactct ttgtcacaag gaaaagagac
46081  cttgtgacac tgaaatgact cgggtgtgtg ggggaacaag ccagcccttt ccctgaagga
46141  ggccctctaa ctctcctctc aggtcctcag ctattaactg gaggaaacag ctgctttttc
46201  agtgcctctc agctactctg tttagctgag agatgaagta ggaagatttg gacttctctt
46261  attgaaaggc ctagagaagg ttttgtgtc cttttaagat gtcacagaaa atttttgttt
46321  caggattgta gggagcagat tcctactgtt cttaaaggac agtaatgcct tttgagtctg
46381  gtctgaagaa cataacaggt ctgtgatcag aagtaggttg catctctctc aactttaaYt
46441  tccttagcta tacctgtagg gatgacttaa gcctagggga gctcctatat ttgggaagct
46501  tgtgcacagg gaagccttaa atgatggtgc ctgcagattg gatctagtag aaattaggtc
46561  cttgggcatg gatgcttggg gaacctctca gtgacctcag gtgaacttgt tgctcgtaga
46621  gccaaagggc gaagttaatt caggccttcc ttttgaccac tgccccctct tcctaggcct
46681  tggccccctc accagaggaa ggtgctgcca cgtgtctgct ccttctgaac ctccaggttt
46741  ctgctacgtt gcccatgga ggacacaccc cctcactca gctgctccga ctgtcagcgc
46801  cactttccca gcctcccaga gctctctcgg caccgagaac tgctccatcc atctcccaac
46861  caggacagtg aggaggtgga cagcatccct cggccctacc gttgtcagca gtgtgggagg
46921  ggctaccgtc accccgggag cctggttaac catcgtcggc cccacgagac tggccttttc
46981  ccctgtacca cctgtggcaa ggaactctcc aatccccatg ctctcaagag ccatatgagg
47041  acacatgctc ctgagggcgg ccgcaggcac agggccccac gccccaagga agccactcca
47101  cacctccagg gtgagacggt gtccactgac tcctggggcc aaaggcttgg ctctagtga
47161  ggctgggaaa accagacaaa acatacagaa gagacacctg actgtgaatc tgtacctgac
47221  cccagcccgag cttcgggtac gtgggaagat ctgcccacca gacaaagaga aggcctggca
47281  agccaccagc tcctgagga tggtcagac ggcctggggac cctccactaa cctctgaga
47341  gcccctctc tccccatccc agccagcagc cttcttagca acttggaaac gtatctggct
47401  gaatcagtag tgaacttcac agggggccag gagcccaccc agtccccctc tgctgaRgag
47461  gagcggcggt acaaatgtag tcagtgtggc aagacctaca agcacgcccg gagcctcacc
47521  aaccaccgcc agagccacac gctgggcac taccctgtg ccatctgttt caaggagtcc
47581  tctaactca tggctctgaa gaaccactct cgaactgcat cccagtatcg gccttaccac
47641  tgtccccact gcccccggt cttccggctc ccccgggagc tgctggaaca ccagcagtc
47701  catgagggtg aaaggcagga gccacgctgg gaggRgaaag ggatgccac caccaatggg
47761  cacacagatg agagcagcca ggaccagctc cccagtgcac agatgctgaa tggctctgcg

```


47821	gagctcagca	cctctgggga	gctggaggac	agtggcctgg	aggaataaccg	gcctttccgc
47881	tgtggggact	gtggccgtac	ttaccgccat	gctgggagcc	tcatcaacca	tcgaaagagc
47941	caccagacag	gtgtctaccc	ctgtctactc	tgttctaagc	agctgttcaa	tgcggtgcc
48001	ctcaaaaacc	atgtgcgggc	tcatacacagg	cccaggcaag	gagttgggga	aaatgggcag
48061	ccatcagtc	caccagctcc	cctgctgctg	gctgagacca	cccacaaaga	ggaagaggac
48121	cccaccacca	ccctggacca	tcggccctat	aagtgcagtg	agtgtggtcg	tgcttaccgc
48181	caccggggga	gcctggtgaa	ccatcgccac	agccatcgga	ctggagagta	ccagtgtctca
48241	ctctgtcccc	gcaagtaccc	caatctcatg	gccctgcgca	accacgtgcg	ggtacattgc
48301	aaggctgctc	gccgaagtgc	agacatcggg	gctgagggtg	ccccagcca	cctcaaggta
48361	gaactcccgc	ctgacccagt	ggaggcagag	gcagccccgc	acacagatca	ggaccatgtg
48421	tgcaaacacca	aagaagaggc	cacggacatc	accccagcag	cagacaagac	agcagacat
48481	atctgtagca	tctgtgggct	gctctttgaa	gacgctgaga	gccttgaacy	tcattggcctg
48541	actcatgggg	caggggaaaa	ggaaaatagc	agaacagaga	ccacaatgtc	acctcctagg
48601	gcctttgcct	gccgagactg	tgaaagagc	tatcgccact	caggcagcct	tatcaaccac
48661	aggcagaccc	accagacagg	agacttcagt	tgtggggcct	gtgccaagca	cttccacacc
48721	atggctgcca	tgaagaacca	cttgcccg	cacagtcggc	ggcgagcag	gcggtactcg
48781	aagcgggctg	gcggtgccag	cggtgggaga	gaagccaaac	tcctggcagc	ggagagctgg
48841	acccggggagc	tagaagacaa	tgaaggcctg	gagtctcccc	aagacccttc	aggggaaagt
48901	cctcatgggg	ctgaaggcaa	cctggaaagt	gatggggact	gtttgcaggc	tgaatctgaa
48961	ggggacaaat	gtgggcttga	gagggatgag	acccatttcc	aggggtgataa	agagagcgga
49021	ggcactgggg	aaggactgga	aagaaaggt	gccagtttac	ttgacaactt	ggacatccca
49081	tgtgaggaag	gtggtggcac	tcacttctgc	gatagcctca	ctggggtgga	tgaagaccag
49141	aagccagcca	ctggccaacc	caactcctct	tcccactctg	ccaatgctgt	cactggctgg
49201	caggctgggg	ccgctcacac	atgctctgac	tgtgggcatt	ctttccccc	tgccactggc
49261	ctgctgagcc	acaggccctg	ccaccaccca	ggcatctatc	agtgtccct	ctgcccgaag
49321	gagtttgact	ctctgcctgc	cctccgcagc	cacttcagca	accataggcc	tggggagggc
49381	acctcagcac	agcctttcct	ctgctgcctc	tgtggcatga	tcttccctgg	gcggtgctgc
49441	tacaggcttc	accggcgcca	ggcccacagc	tcctctggca	tgactgaggg	ctcagaggag
49501	gagggggaag	aggaaggagt	ggcagaggca	gccccctgcac	gcagtcacc	actgcagctc
49561	tcggaagcag	agctgctgaa	tcagctgcag	cgggaggtgg	aagcgtgga	cagtgcaggg
49621	tatgggcaca	tctgtggctg	ctgtgttcag	acctacgatg	acctggggag	ctcggagcgt
49681	caccacaaaa	gtcagagtcc	tgggactact	gcagacaagg	ctcccagccc	cttgggagtg
49741	gcaggtgatg	ccatggagat	ggtcgtggac	agtgtcttgg	aggacatagt	gaattctgtc
49801	tctggagagg	gtggagatgc	caagtctcaa	gagggagcag	gcacccctt	gggagacagc
49861	ctctgcatcc	aggggtggga	aagtttggtg	gaggtcagc	cccgccctt	ccgtgcacac
49921	cagtgtggca	agacctatcg	ccatgggggc	agcctggtga	accaccgcaa	gatccaccag
49981	actggagact	ttctctgccc	tgtctgctcc	cgctgctacc	ccaacctggc	tgccatccgt
50041	aatcatctgc	ggaaccaccc	tcgctgcaaa	ggctctgagc	cccagggttg	gcccattcca
50101	gaggcagcag	gtagcagtga	gctgcagggt	gggccatcc	cagaaggagg	cagcaacaag
50161	ccccagcaca	tggcagagga	ggggccgggg	caagcagaag	tcgagaagct	ccaggaagaa
50221	cttaaagtgg	agcccctgga	ggaagtggcc	aggggtgaa	aagaggtgtg	ggaggagacc
50281	actgtgaagg	gggaggagat	agagcccagg	ctggagacYg	ccgagaaggg	ctgccagact
50341	gaagccagct	ctgagcggcc	cttcagctgc	gaggtgtgtg	gccgatccta	caagcacgcc
50401	ggcagcctca	tcaaccaccg	gcagagccac	cagaccggcc	actttggctg	tcaggcctgc
50461	tccaagggct	tctcaaacct	catgtccctc	aagaaccacc	ggcgcatcca	tgcatatccc
50521	cgacgtttcc	gctgcagcga	gtgtgggaag	gccttccgcc	tgcgaaaaca	gctggccagc
50581	caccagcggg	tccacatgga	acggcggtgg	ggtgggggca	cccgaaggcc	gactcgggaa
50641	gatcggccct	tcgctgttgg	gcagtcgggg	cggacctatc	gccacgccgg	cagcctcctg
50701	aaccaccRgc	gcagccacga	gacgggccag	tacagctgcc	ccacctgccc	caagacctac
50761	tccaaccgca	tgggcctgaa	ggaccaccag	aggctgcact	cagagaatcg	gcggcagcgg
50821	gctggacggt	ccaggcgcac	agctgtgcgt	tgcgccctct	gtggccgag	cttccctggc
50881	cggggatctt	tggagcggca	cctgcgggag	catgaggaga	cagaaaggga	gccagccaat
50941	ggccagggag	gcctggatgg	cacagcggcc	agtgaggcga	acctgactgg	cagccaggga
51001	ctagagaccc	aattgggtgg	tgctgagcca	gtacccact	tggaggatgg	agtcccaagg
51061	ccaggggagc	gcagtcagag	ccccatcagg	gcagcaagct	cagaagcccc	agagccactg
51121	tctgggggtg	caggggaaggc	agtggtgtg	ccggtaggtg	ggggactggg	gaatcactat
51181	ggagRctggg	ttcctcagtt	cctaactagg	tcagaggagc	cagaggacag	tgtccacagg
51241	agtcccttgc	acgctggtga	ctgccagctc	aatggacct	ctctgagtca	catggatagc
51301	tgggacaaca	gagacaacag	ctctcagctg	cagccaggga	gccactcctc	ttgcagccag
51361	tgtggcaaga	cttactggca	gtcaggcagc	ctcttgaaac	acaacaccaa	caagacagac
51421	cgacactatt	gctgtctctg	ctccaaggag	ttcttaaatc	ctgtggccac	aaagagccac
51481	agccacaacc	acatagacgc	ccagaccttt	gctgtcctg	actgtggcaa	agcctttgag
51541	tcccaccagg	aactggccag	ccacctgcag	gctcatgccc	ggggccacag	ccaggtgcca
51601	gcccagatgg	aggaggccag	agatcccaaa	gcccgggactg	gggaggacca	ggtggttctc

51661	cctgggtcaag	ggaaagccca	ggaggcccca	tcagaaaccc	ccagaggccc	aggagagagt
51721	gtggagagag	ccaggggagg	acaagcggtg	acgtccatgg	cggctgagga	caaggagcgg
51781	cccttccgct	gcacccagtg	cgggcgctcc	taccgccatg	ctggcagcct	gctgaaccac
51841	cagaaggccc	acaccacagg	gttgtacccg	tgctccctct	gtcccaaact	tctccctaac
51901	ctgctgtctc	ttaagaacca	cagcaggacc	cacacggacc	ccaagcgcca	ctgctgcagc
51961	atctgtggca	aggcctttcg	gacagctgcc	cggctggagg	gccacgggcg	ggtccatgca
52021	ccccgggagg	ggcctttcac	ctgcccccat	tgtcccgcc	acttccgccc	ccgaatcagc
52081	ttcgtgcagc	accagcagca	gcaccaggag	gagtggacgg	tggccggctc	cggtaggggg
52141	catgaagggt	cccaggagga	ggtgggcaca	cagtggaggg	ggaagtccag	cccccaggtc
52201	ggtgggggag	caaggagtga	gaggagagag	ccccggggat	tctaagaggt	gggtgggggc
52261	ttggctatgg	ggtgagagaa	gtagcttgag	gatgtgctga	gctgagcacc	cgcaagtcag
52321	gtataacaaa	tagcagggtg	ggttgggcag	cacgtggggg	cgtggtcagg	ccgaggctgc
52381	tacctgggct	cctccattac	actgtagcca	gaatggaatg	gtctttctgt	tcaggggaag
52441	gtcactgggt	accccctggc	tgctgtgtct	ggaaaccctc	ctgagtcagc	cagtaaagta
52501	atgacttcca	gagaaaaaga	ggaagccatt	ggtttggtct	aggttccatt	ctttcctgga
52561	gcaggccggg	tgccagggaa	caagggatgg	ggcatgggct	ccacggcttc	cctgtgact
52621	tggccacgga	aactggttca	ctggttggca	ccctactccc	tgtccctctt	tccctgcgcc
52681	ttgtctctgc	tgctcctctc	cttgaaact	agacctctgg	tccttccctg	tcagtgttgc
52741	tcccatctct	tctctaacct	ttattcagcc	ccttttccct	ctgctgccaa	cggccttttt
52801	aggatccaac	caaaccaccc	tttctacctg	cgcaccctgc	caccctctgc	acaccttaa
52861	ctggagagct	gagtcacaga	taattgtttc	cttgaagtcc	aggccagctc	gcacaacaa
52921	cagtcattag	cccgtgtcac	atccctgatc	agagggcatc	tccgtgggga	atcgctcca
52981	cccagcactg	ctggaagccg	cRgtgccag	ggagtggggc	ggccgggttc	ctcagcagga
53041	cctgggctgg	cctctccacc	tccYctagta	gaggcggacc	cattccatct	agtggccacc
53101	gagggtgggt	ggccctgaga	tggtggggcc	ttgacaggcc	ttgtcagagc	agagggcagg
53161	tggtgagtcac	ctgaaagctg	aaggaaatgg	tttaaggata	gaagatttct	catgacctca
53221	agggatatga	gggaggagcc	agtttgccag	ggctgggaaa	ataattagga	ggcctagaat
53281	ccctgttctc	atctgggcct	ccggggccag	gggcagggga	atggcctgca	gggctgggag
53341	ggggtacacg	ctgtgcgggg	tctgcccctc	agttggtgac	ctcctctctc	tctcccccca
53401	ggagccccag	tggcaccagt	gacgggcaga	ggggacttgc	cattgcccc	tccaccacc
53461	cccacgaccc	cactcctgga	tccttcaccc	cagtggcctg	cagacctcag	cttctccctc
53521	tgaacttcaa	gtctccaaag	atcagaatct	gggggaggga	gcgctgagc	ggaggggctt
53581	gatctccaca	ttttctcagg	agtagttcgg	gcatcccat	atcttctcct	ctccccYtgt
53641	gaagaggacc	cagatctggc	ttctttccca	aggagggggt	ggggtgttcc	tcgctccct
53701	gtccttgaag	gacctccttc	ccccagcctc	atcacctgct	tcttctcagc	gccacctca
53761	gcagccagat	tgcaacacca	gggagaggcg	gatgcagagc	cccaccgggt	ggaaagttgc
53821	ctgtggaagg	gagccttttg	ctacaatttg	taacttattt	tctaaagtct	attttgtaac
53881	aatttattta	agtttaaaaa	aaggaaaact	gctgcccccc	aaaaaaagaa	attttcaaaa
53941	caacgtggct	ggcgtgattg	tatctgaaag	ggtaaaggag	gaggaaagct	gagacgcctg
54001	cttggtagca	gagttgggtg	tgggagtgct	cacagacacc	cctgtcctgc	aggggtggga
54061	gtgggcacct	gtggcccccag	gcaggttctc	tcccacagct	cgtgggcttc	tgggctggcc
54121	ctgggtgctg	gaatcacaca	tgacagggtg	gggaggacag	gggcagtaat	gccatttgcc
54181	tgcttgcat	ctcttgctct	gagaatggcc	aggtcccctg	tcagcagctg	ggtgggtggc
54241	ctgtggggaa	ggaaggaggg	tggagtgtgc	ctcatcctca	cggctttggt	ccctccctcc
54301	ctccccattc	ctcgaaggaa	cagggtctgt	cttggccgct	atgacagatg	agaatactga
54361	ggctcaaagc	ggttgagcag	cctgtcccaa	gtcacacgat	gacaaagaac	cagaatctga
54421	atcaaattgg	tctgcctgtt	gtccaccct	acccaaggca	gctggagtgg	gttagaacgg
54481	cacgttctca	ctggagagag	aagggtcctg	gagaggcagg	gtttggcagg	aggccccggg
54541	gccacatact	tatgttggcc	aggcagcttc	caggctcagc	ctcgggctct	ggttcctcgg
54601	cgaagtagac	ctgccagtc	aaactgctga	cccagtcctc	ataggcaggg	agcgcggtga
54661	agaccgccc	cctggcgggg	ccttggcaag	catctccgaa	gctgtgcagc	ccggccagga
54721	accatgtgcc	cctcacctca	tgaccagtg	gtgccccaga	caggccctgt	caggggtcag
54781	gtgacactgg	gtgacttttt	ataggcagct	gtgcaggacg	gtagagcagg	ctagggaaac
54841	tctggctgtt	tgggggctaa	ttggcaaaaa	ggcttagttt	caggtgggag	gtgggtccag
54901	aggccaagcc	tgggaaggtc	gttcttgagc	tattgggtga	ggggtatgca	ggggcctgtt
54961	aggtgaggtg	gtggggcctg	gggtccacct	cacagctggg	cagctcacc	acagcactgg
55021	tacacaccat	ccccggcaga	atagggtgct	catcaccccc	aggagctgca	tgcagccggc
55081	tgaggccct	aggccccagg	agggtcacgg	gcactgtctg	gagggagctg	atgcctgtgg
55141	agcaagggaa	agctggctgc	cccgccctgc	aggttggatg	gacagcagcc	ctggccctgt
55201	gcccaccctac	ctgctcctgg	gcgggcccgt	cccagaaccc	agccacgctc	cccatcaggc
55261	aggtgggtgt	cagSataggg	caggccaggg	ggccgcaggc	tggctccag	tgccacaggc
55321	tgggcccagca	gcaggagggc	catgtcgtag	ccccctcag	ggtgggtgta	ggctccatgc
55381	aggatgagct	gcttcaggcc	ccactcctcc	ggtctggtcc	ccagccctac	gctccattcc
55441	tctggggcct	ggcgccctgt	cagaggcag	gtggacggca	ccaggtgaca	gggctgccgg

55501	caggggagcag	ggggcacagca	gagaggggac	caagactcac	ccaatgaagc	agtgggagcag
55561	agtttagcacc	gcctcctctg	acaccagggc	tccgccacag	gccagctgtc	cctgggtgcat
55621	cagcctggcc	tcccagggcc	atggggaggg	tgctcctgcc	tggggacctg	ctgtcctcaa
55681	ggatccacag	gctgaaacag	ataagtgcct	catctgactt	cttgctaagc	acttcccact
55741	tcccatcaaa	tcctcacagt	agcttttgaa	attgaatttt	gttccccca	ttactgaagg
55801	taaccaaagt	tccagagagg	ttaaRagact	tccttgctct	gtaactacag	gcaagtggca
55861	gacttgcat	tcgcaagcca	gggcctgagc	ttttagccct	gttggtctga	tgggaccct
55921	gtcccagcct	cgtgggaatg	cccctacttg	tcactccct	gctccaccg	tctgcaccct
55981	ccaggMtggg	cctgcctccc	gtattcctga	ctcattagt	catctcttat	catcaagaca
56041	taatatccgt	cccaggaatt	actttgcatt	caatttgcat	gccagtcact	taatgccgga
56101	attctgactg	ggagccctac	cctgtgcagg	ctcgtctggg	ccctgctgga	agcctgccac
56161	ttccccagaa	acccaagtca	ggtctcagag	attcctcttc	tcacctaaac	tccaaacctg
56221	tagagttcca	aagtgcctgt	gcctctcagc	cctaacaggg	ctgttcccat	cccagggggg
56281	gaaagagccc	cctaattagg	ctcggcggtg	tggatgccta	tgccagttct	ctagtcctaa
56341	ctgaggtttg	cttcacagt	gcttctgccc	actcccagca	tgccccactc	gtacctacac
56401	agctgtcctc	atcactcatc	tcgggggtct	ctgggctctg	ggccaggaaa	gctgccccct
56461	gaactcgagc	ctgcagccag	gaactgtgag	cagctgtgtt	ggtcagcagc	acaggagcgt
56521	cctcctgggc	acagcttgat	gcaaagctga	tgatgccagc	ctgaaccag	tgtccgtcag
56581	gctcgaggca	cagcacaggg	cccccggaat	ctccctggag	ccaggcaaca	aagccaagga
56641	catagcctg	agcccagcta	tggcagacac	cctctgattg	cagStctttc	ccccagtcct
56701	gtgagtccaa	ccccagccc	agggttaggc	ccctcccttc	ctgtctcttc	tcttctccct
56761	atcagacctg	acagggggcc	tgcaccccc	gctggggggc	cccacatagc	atcccaggcc
56821	gggcccgggt	ggacaggtgt	cgctgggtga	gctgggtgta	gatacagtta	catgtggggc
56881	gactgatgag	acgcaggcgc	agattgctga	gggtcccagg	agctggtgaa	agagacgggg
56941	ctggggctag	agtctgggat	ctgggaagca	agggagactg	gaagccagga	cagttggggg
57001	gtgagactg	gacgtgggc	aagcagagt	cctctccgag	cttcagtctg	ggccagccat
57061	taccatcact	gggtgtcctga	tcccagccag	tggcccagca	ggaggctcca	aaggggaagc
57121	gatgggccc	ctggggcagg	cagaggggtg	tgtgggtcgt	gggggtggcg	agctgcagca
57181	gggccaggtc	tgagccctgg	ctgtagtgtt	tataggccct	gggcaactgc	agggcagcca
57241	ccccacctc	ttcggcccca	gggctgagtc	cctcacgctg	cagagaaccc	aggaccactg
57301	accaggaatt	cagttctggt	gctgtgccc	tggaaaggga	aggacaaggt	ccaggctgtc
57361	gagtgaagg	gtagacagt	acaccatggg	agaccccaca	ggtaggtgga	agatagcaga
57421	gagcactcca	cggctttttc	tttttgaggt	aggtctctgt	cctgtcggcc	aggctgcagt
57481	gcagtggcat	gaccatgact	cactacaagc	cttgacctcc	caggctcaag	cgatcctccc
57541	acctcagcct	cgtgagtage	tgggactaca	ggcgtgcacc	accatgcctg	gctaatttta
57601	cagtttttgt	agagacatgg	tcttgctatg	ttgcacaggc	tagtcttgaa	ctYttgggSt
57661	caagcgatcc	tcctccgtg	gcctcccaaa	gKgctgggat	tacaagcgtg	agctactgtg
57721	cccggcctca	gtctgtgact	ttaagcgagg	tcattccctt	gcaacctgag	ttttYtgatt
57781	tgtaaattct	ctaccatgaa	aggtgggtgt	tacaattaac	ttgacaatat	gtacccagta
57841	catggcagat	tctcagtaaa	tggttcttta	ttattactct	agaggcaggg	actagccttg
57901	tggggcttga	gacaaattgt	ccactcccca	gagccggaga	ccctgcctga	ccccagccat
57961	gacttacttt	tcaaagcagt	gggcagcagt	gaggaccag	gtgtctgcca	ccaggagacc
58021	gctgcagatg	tgggctcctt	gcctcctcac	actggcctgc	cagggccact	cgccagggac
58081	tgtgttgccc	tcctgaggct	tggggggggc	ggggccacgc	tgtccacagg	ctgtggaaag
58141	gagttagtca	cactgacagc	agatacctgt	cctgacctca	ccccagtcct	aaccagacc
58201	caagtaagac	ctaggccccg	tgtccccaga	ccttaccacg	ctgagcggct	tgaagacctg
58261	ggaagagaga	gacacaggta	agatgcaggg	actccaggcc	tgccatgctt	tggggaggag
58321	aggggatggg	ctggggggcg	ggcccagggt	ccttgccctg	gactctgatt	accttaccag
58381	cccctcccag	aatgaaaata	tttatatgag	gccgagacag	cttttattgg	aacctattca
58441	gtgtgcacac	tcagtaatta	attctccttc	agctgtgctc	agcactctgg	ggcttggggg
58501	tcagcttggg	agtaggccag	ccctcctcca	ggcttcagaa	cccccaactc	ctgcccccg
58561	cactgagtca	gccaggcggc	ctgtgtgtgt	agagagcatt	agcttaattg	tcctcttagc
58621	agcagaggcc	taagaggaag	gatttagagg	ctgcatcatt	tccaagtggg	gagggcccca
58681	agaaatggag	acttacctca	ccctgggtct	gagactcagt	cttccccacc	ttcccagaaa
58741	ctgtctgaga	gcccggccaga	gagagggccc	ctgcccaccg	cccctcacag	gcacacaggc
58801	acccccagag	acagctgagc	caggctgccc	agaggatgga	tgaaaagaaa	gggaaactga
58861	ggccagagga	gcccagagtt	tgccctgacgt	cattgtggaa	gtcgaggggg	aggcaggcac
58921	aggacacacg	caggcagcac	ctcacacaca	cacaagacca	caggccccgc	caacgcaaac
58981	tgacgtggc	ccgagaaaa	ctcatccatg	ttgacacagg	tggccacata	taccaccca
59041	cagagtccta	cggaagttaca	tccccactgg	tggactgtcg	cccacaggcg	gtccccccac
59101	caagaagcag	ggactgctgg	ggcagagatg	gcccctgagc	ccccaccagg	ccacacccat
59161	acccccagcac	atggcggtct	accctccatg	aggactgtgg	cacccgcgat	gagcagcact
59221	gggccccagc	accacttcat	gctgccccgg	gccactctgc	cacctgtgct	ccactctgag
59281	agaggccacc	tgggtctccc	tggctccacc	tctgtctcac	ctccagttgg	ctaggattca

59341	gctgtctgct	tgccctagcc	agcagttcct	agctctgggg	ctgtggctgt	gtgaccctag
59401	gcttgtcacc	acccctctct	gggccttagc	cgtggagcca	aattgtcttg	gtttaaatct
59461	gtttccttac	tacttgagtg	gcctgggaga	gctaacctgc	ctgtgccttg	tctgagaaac
59521	agttgttgaa	tggaacacgt	ctcaatgaca	gcctggcaca	gggcaagctc	actcccagca
59581	tgaaagtgcc	atcctggaac	catggagggc	acgtgggcag	ctcctcccag	tccccacttt
59641	tcatttccac	tccccactt	ctcttccagc	cagggctagg	tggttttcac	atgccttctc
59701	ccagggtctgc	tcagagacag	tgtgtgccag	gctgagtggg	gaggagccag	caccactgct
59761	tgcttgctcg	ctgtgagcct	ttggcaaadc	ccagtatctc	cctgggcctg	ctggcttctc
59821	gggaaacagg	gtgacagcgg	tgcccacctt	cctcataggc	gcagaggact	cggtcaggcg
59881	ggaagaacac	tttggcgcg	gctcctctct	cctcttccca	ctcctaactc	aggctggccc
59941	agaccattca	agaacctga	cccaagaca	gaggcagctg	tggttaagg	ttagcatata
60001	ttatagatgc	tgcccgggca	tggtggctca	cgcctgtaac	cccagcactt	tggtgaggcca
60061	aagcaggcag	atcacctgag	gtcgggagtg	cgagaccagc	ctgaccaaca	tagagaaaca
60121	ctgtctctac	taaaaaatca	aaattagcca	ggcctggtgg	cgcatgcctg	taatcccagc
60181	tactcgggag	gctgaggcag	gagaatcact	tgaacctggg	agggtggagg	tgtggtgagc
60241	cgaaattgtg	ccattgcact	ccagcctgga	caacaagagc	gaaagtccgt	ctcaaaaaca
60301	aaccaacaaa	aaaaaaagca	tatagatgct	ggagccagat	ggaccagggt	gattcaaatc
60361	ctgagagcct	cctgcagtag	ctgtgtgatc	tcaggcatac	taattagcca	ctctgtccct
60421	atcatcaaga	tagggataat	aatagtaccc	acctcataaa	tcactgtaaa	gattaaacga
60481	gtttaacacg	ttaacacaag	tttttttttg	ttttttgttt	ttttttggag	acggctcttg
60541	tctgttgctc	aggctggagt	gcagtgatca	tggtcactg	cagcctcaac	cttctgggct
60601	caagcagctc	tcccgcagc	ttcccagata	gttgggacta	taggtacata	caccaccatg
60661	cctgggcta	tttttttttt	tttttttttt	gtagagacag	gggtctcact	attttgccca
60721	ggctggctctg	gaactcctgg	cctcaagtga	tcctcctgcc	ttggcctccc	aaagtacagg
60781	gattacagg	gtgagccact	gcaccctgcc	aacacaagtt	cttaaacagt	gcttggcatg
60841	taggtaagtg	gtcagggcat	aataggcaaa	acaaaaacct	tcacaacctg	gcccagacc
60901	tccaaggcta	ccaatactgg	cttgggaatg	tcgataaggc	aactggagg	gttaaagtta
60961	aactcaagga	agaacttccc	agcaagcatt	tacagaacca	gaagggcagc	ctgcccttcc
61021	agggtgtgtg	tcagccttcc	caggagagga	ggcctggctc	ctgtgggcag	caggagcgag
61081	ctgcagcct	gtttcctggg	ggtggggggc	aatggtgccc	caggccttgc	tgactccaca
61141	cactggagat	gagactacc	ataaccaccc	ttcccagcag	gccctccact	ctccctctga
61201	ctcacccttc	ccagctccag	agaaggcaac	accgagggag	gccagcacc	acagtccatg
61261	gcagacacat	ggttcagact	tggtgattg	atctaagaaa	ctttattgct	caRaaccttc
61321	cctccctggg	caatggaaag	agctttggag	accagcccat	ggggacagag	tcagaggcac
61381	tggtgtgtaa	aaagagcgag	cgtgtggcac	atgtgttcca	ttgtcatgtg	Ygggtatggc
61441	aggaggagg	ggtaatctag	aagcccaca	tctaggccct	tctagggacc	cagatatgcc
61501	cccttaggca	aggctcacat	gccaaagcaa	agcagatgag	gtcagcctgg	cttgggttga
61561	gggtcagtg	cctcttagcc	ttgccctggg	gttcttggac	cttccgaaa	ctgagccaca
61621	tcaggctcac	gttgatagca	taggtgtgga	tacaaacaat	gcagaaatca	tagagcacga
61681	agaacaggat	ccaggccaRg	tagacagaac	cagcgagaga	caccagggag	ctcagcagca
61741	tcaggacaga	ggcccagcgt	gtccgcaggg	aacctgcaag	gcagaagagg	gtccggtgtg
61801	ggcttcaggc	actggccacc	tcccgaacac	tccatgatgt	cactgcacta	cctagatgcc
61861	aggagctgct	gggaagggtc	tctaaaacaa	gaggctccag	ggcagatgtg	gtggcttacg
61921	cacgtattcc	aaacactctg	ggaggccgag	gtgggaagac	tgcttttagc	taggagttca
61981	agaccagctt	gggcaacata	gaaagacccg	atctctatca	aaaatttaga	aattcagctg
62041	ggcacgggtg	ctcatgcctg	taatcccagt	attttgggag	gccaatggg	gtggatcacc
62101	tgagggttagg	agttcgagag	cagcctggcc	aacatggcaa	aaccccatct	ctaataaaaa
62161	tacaaaaatt	agccgggtgt	gttgatgggc	acctgtaate	ccagctgctg	gggaggctaa
62221	ggtggagaat	cgcttgaacc	taggaggtgg	aggttgcagt	gagccgagat	catgccacca
62281	cattccagct	tgggcagcaa	gagcaagact	tcgtctcaaa	aaaaaaatgc	ccggtgaggt
62341	tgactcacgc	ctgtaatcct	agcacttttg	gaggccaagg	cggtatggtc	accaggtcaa
62401	gagattgaga	ccatcctagc	caacatggtg	aaaccctgtc	tccactaaaa	atacaaaaaat
62461	tagctgggag	tggtgacacg	cgctgtagt	cccagctact	caggaggctg	aggcaggaaa
62521	attgcttgaa	cccgggaggc	agaggttgca	gtgagccaag	atcgtgccac	tgactccag
62581	cctggtgaca	gagcgagact	atgtcgcaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaatttagc
62641	tggtgagctg	gacatgccta	tagtccttagc	tactcaggag	gatcgcttga	gccaggttac
62701	agtgagctat	gattgtgcca	ctgcactcca	gcctgggcaa	cagagcaaga	ttatttctta
62761	aaaaaaaaaa	aaaagaaggc	ttcaacagg	cccctccaag	ggactggtct	ctgaagctct
62821	tgccattgcc	cagggaggga	aagttctgag	caataaaatt	tcttaataaa	atcgccaag
62881	tctgaaccat	gtgtcagcca	ggaccRtggt	gctgggcctc	cctcagtgat	gccttctctg
62941	gagctgggaa	gggtgactca	aagggagcgt	gggagcctgc	tggaagggtg	ggtaattgat
63001	agtctcatct	ccggcatatg	gcacagcaaa	ggcctggggc	gccatcgtct	tccactccct
63061	tggttcctct	ctctgttctt	atgggactag	atacaaaatt	tcttgctgag	cactaaatga
63121	gacaaaagat	agctcatgct	cagcttctcc	ttaaaaagga	atctcgcat	cttttccaca

63181	aaactggggg	gttggtgggg	catggtagct	cacgcctgta	atccccccag	cactttggga
63241	ggctgaggca	gacagattgc	ttgagaccag	cctgggcaac	atggcgagac	accatctcta
63301	ccaaaaaaa	acaaaaacaa	aaattagctg	ggcatagtgg	tgcacgcctg	tgattccagc
63361	tgcttgggag	gctaagggtg	gaggatccct	tgggcagggg	ggcagaggtt	gccatgaact
63421	gagatcacgc	cagtgcacac	taagggcac	ctagacctca	ctttgggcaa	cagagccaga
63481	ccctgtctca	aaacaacaac	aaacaaaaaa	cctggggacc	taggatgtct	ttaagggtcc
63541	ttcagcctct	aacagtactt	aaaccaatta	aaagactcct	gttagttacc	tccccacatc
63601	cccacccgca	ggacgctcSg	tgatgagcag	ctagctggct	gtcagctgtg	tggatcacca
63661	agattgcatg	gagtggggct	gagctgacca	agggggatga	ggggcggggc	ggggcgggca
63721	gggagggggc	ggagccactc	acctaacaat	agctgtagt	tgtagaagat	gcaaccgaat
63781	atgctgttgg	attgattgag	gatgctgtcc	tgtcccagca	catgctccac	cagcccga
63841	cccctgcccc	acctggcaga	ggggtggggg	ggggtggaac	caggttagga	ctgtcaaccc
63901	agtgccttgg	accctgcccc	agaaagggtga	tttccaagaa	gccacctggg	ctatcctctg
63961	ttccccgacc	tcccatccta	gtccaagggt	cgatgatctc	ctggcaccgg	gcacctttgg
64021	ccacgtcagg	attccatgtc	actgacccta	tcctcccctc	tccccagacc	aggcccggac
64081	gtggctactc	cgtaggccct	gcttttcac	ttagacctta	agtaagtctc	tttttttttt
64141	tttttttttt	tttttttttt	acggagctct	actctggccc	aggctggagt	gcagtggcgc
64201	gatctcggct	cactgaaacc	tccgcctccc	gggttcaagc	gattctcctg	ccttagcccc
64261	cccgagttagc	tgggattaag	gcaccgcgca	tagcgcccga	ttaatttttc	tatttttggg
64321	agagacaggc	tttccaccatg	ttggccaggc	tggccaactc	ctgacttcaa	gtgactccatc
64381	cgactcgccc	tcccaaagt	ctgggattac	aggagtgggc	caccgcgccc	ggccttaag
64441	taattcttaa	aatggcaagg	ctggtataac	gggtcactcg	gttttgcac	agagactggg
64501	agtcgggggc	agattatctt	tgccttgga	cccagaatct	ccagctccct	ggccactcac
64561	tcgcctcctc	tgtattccgt	cattatgcta	acgcctggcc	atcacgcaca	gccagaccgg
64621	gccaccttgt	tcctgggcgc	agccatcgcc	aacaccccc	ttcMctgcg	cgccgtcctt
64681	gagaccatcg	tcaatctcta	ccgcctcct	gcctcccgc	ctttcctggt	cacccttatt
64741	ccttggcatc	caattcacgt	gcgagtcccc	ggaataatcc	cagtccccag	cactgtctgg
64801	tcccttgcc	cgactctta	tttcgaacac	cagtatcgct	gggtagctca	gcccctgtgc
64861	aacgaccccc	cgagcagtc	agccccgtgt	ccgttccccg	ggcacaccga	tcccagactc
64921	cagaataatc	atctggcatc	ctggccgccc	tgtccgagg	ccccacgcct	cccactcccg
64981	tgcacacctg	gaggagaaga	cgcgcgaaca	gctgatggcg	gtgcccacgt	cgagagcg
65041	gcggtaatcc	cggtccccgg	cgcgcgcgc	cttcacgtgc	agcgcgtaga	gcgagagcac
65101	taagcccgtc	aggcaaagag	cgagccgcac	ccagccaggg	ctccccagg	tgtgcccac
65161	tatctccagg	ttccgcccga	ggcgcccgcg	gagaaaacca	gccacggaac	aggggcccgg
65221	cggcgaatgg	ccgcgcccct	cctggcccct	tgactcgggc	attggccggc	cgtgctcgca
65281	ctccacgacc	caaatggctg	ttccagggcg	ctagtcaagc	ggcgaggtta	ggaaaaacag
65341	gaagaatgcc	gggactagt	aagcgggtaa	gggacgtgcg	gaatcgcggc	cccagcggct
65401	gccaggcatg	atgggagttg	tagtcggcg	ggctgcaagg	catcaaggga	aatgaagtct
65461	ccacagattt	aaaaactggt	ggcggggcac	ggtggtcac	gcttgtaatc	ccagcacttt
65521	gggaggccga	ggctggcgga	ttacctgagg	tcaggagttc	aagaccagcc	tcgccaacat
65581	ggtgaaaccc	catctcaact	aaaaatacaa	aaattagccg	ggcgtggtgg	cacatgtctg
65641	taatcccagc	tactcgggag	gctgaggcag	gagaattgct	tgagccgggg	agccggaggt
65701	tgcagtga	cgagatcg	ccactgcact	ccagccaggc	cgacagagtg	agactctgtc
65761	tcaaaaaaaa	aaaagaaaaa	aaaagtggtg	ttagtgtggt	taacagcatt	tgcgcttacc
65821	ctatgccaa	tcctgttgta	agaattgcag	catccgggac	ctagagacca	gcggatcagg
65881	ggatccagcg	aatacggcga	tccgattcgg	gaaccaagca	tttcccctga	aactatttca
65941	ggcaccattc	gggctgcagc	ctcccatcct	cccgggtcct	gcctcaccag	tgttctctgg
66001	tggctcggtc	ccctttctcc	catacattca	cagaaccact	cctttggcca	cacacaccct
66061	tgacagtatc	ctaacctctc	tactatcatg	gcgcccgcct	ggcaacacct	gagcttgcac
66121	caactcaaca	gtcagtctct	cctttccaag	ctgtgggcca	gatgaggtct	ctccttcaca
66181	gacgtcccat	ctgatggtga	cccatctctc	ctacaccttc	aacctctcca	tcttcccat
66241	ctacactgca	acttgtttct	cttcccac	tcagaaacag	caacaaatct	cccttacta
66301	agaggtcctt	caccagcctc	ctctcccggc	attatcccat	ctacccctcc	acattcaagt
66361	ttttggaaag	attctacact	cccagtctct	acttcctcac	ttcttctctg	ctgcccacgc
66421	cataaactag	ctgctgcctc	cagcattgcc	ctgacacctc	gtggctgggtg	tcaccaagac
66481	gctagaccga	atggttattt	atttatttat	ttacttattt	tgagacggag	tctcactctg
66541	tcgcccaggc	tggagtgcag	cgggtgccatc	tcggctcact	gcaacttccg	cctccagggt
66601	tcaagtgggt	ctcgtgcctc	agcctcccaa	gtagtttgga	ctacaggtgc	ctgccaccat
66661	gtctggctaa	tttttgtatt	tttagtagag	acagggttcc	accatgttgg	ccaggcttgt
66721	cttaaacctc	tgacctcaag	tgatccaccc	acctcgccct	cccaaatgc	taggattata
66781	ggcgtgagcc	accgcacccg	gccaatgggt	gtttttcagg	tcttctcttg	cttgacttcc
66841	cagagggatc	ccttactggt	gcacctaccc	ttctgggaac	tcttctctc	tggcgtctgt
66901	gatatttccc	tctcctgctg	gctcctccct	ctccagatgc	tgtttctcac	atctactctc
66961	ttctagagag	tgtggtagac	agaataatgg	tcaccaaaga	tgtccctgca	tgaatccctg

67021	gaacttgatga	atatgataggg	ttaaattggcc	aaaagggaat	taaggttgca	gatggaatta
67081	agctgaccaa	tctcctgatt	ttatttttatt	ttattttgtt	tttgaggtgg	agtttcgctc
67141	ttgttgccca	actggagtgc	aatggtgtga	tctcggctca	ctgcaacctc	cgctgcccag
67201	gttcgagaga	ttctcctgcc	tcagcctccc	gagtagctgg	gattacaggc	acccgccatc
67261	atgcctggct	aatttttttaa	atttttagta	gagacagggt	ttcgctatat	tggccaggct
67321	ggctctgaac	tcctgacctc	aggtgatccg	cccacctcgg	cctcccaaa	tgctgggatt
67381	acaggcgtga	gccacctgtc	ccagcctatc	tatctattta	tttattttatt	tttgagatgg
67441	agttttgctc	ttgttgccca	ggctagaatg	caatggtgct	atctcgactc	accgcaacct
67501	ccacctcccc	ggttaaagcg	attctcctgc	ctcaggctcc	tgagtagctg	ggattatagg
67561	catgtgccac	cacgcctggc	taattttttg	tatttttagt	agagatgggg	tttctccatg
67621	ttggctcaggc	tggtctcaaa	ctcctgacct	cagatgatcc	accaccttgg	gcctcccaaa
67681	gtgctgggat	tataggcgtg	agccatcata	ccaggctcta	ttgattttatt	tttatttttta
67741	tttttgagac	ggagtctcgc	tctgttgcc	aggctggagt	gcagtggcac	aatcttggct
67801	cattataact	tccgcccccc	cccaggttca	agccattctc	ctgcctcagc	ctcccgagca
67861	gttgggacta	caggcgctgt	caaccatgcc	tggctaattt	ttgtattttt	agtagagacg
67921	gggtttcact	gtgttgccca	ggctggctc	gaactcctga	ctttgtgatc	tgccctgctc
67981	agcctcccaa	agtgtcggga	ttacaagtgt	aagccaccac	gcccagccta	ttttgtttat
68041	tttttcaaa	acccttgaca	cccaggctgg	agtgtagtgg	cactgtcata	actcactgca
68101	acctccgtct	cccaggttca	agcgttctt	gcacctcagc	ctccctagta	gctaggagta
68161	caagtacgtc	ccaccacacc	tggctaattt	atttttattt	ttgtagagat	ggggtctcac
68221	tttgtttccc	aggctggctc	aaacttctgg	tttcaagcaa	ccttcccacc	tcaaagtgct
68281	gggagtacag	gcattgagcca	ccaccacacc	tggcctaatt	tgctgatttt	tattttattt
68341	ttattattta	tcttaatttt	tattttgaga	cagagtcttg	ctctgtcatc	caggctggag
68401	tacggtgggt	caatctcagc	tcactgcaac	ctccccctct	cgggttcaag	cMattcttgt
68461	gcctcaacct	cccagtagc	tgggattata	ggtgctggcc	accacgcctg	actaattatt
68521	tgattttttt	ttttttttag	tagagcggg	atttcacat	gttgggcagg	ctggctttga
68581	actcctgacc	tcaagtgatc	cacctgcctc	agcctctcaa	agtatgggga	ttacgggtgt
68641	gagccgccgt	gcctggccca	atttttgat	tttcagtga	gatgggggtt	tgccatgttg
68701	gccaggctgg	tctggaactc	ctgacctcag	gtgacctgcc	tgccctcgcc	tctcaaagt
68761	ctgggattac	aggcataagc	caccatgcct	ggcccaacag	ggctcttaa	aaatgaagga
68821	ggatggcaga	agaaagtacg	aggagatgt	gagtaaagaa	aaaagacaca	gagagctgca
68881	atgtttcttg	tttgaagatg	gaggaagggg	attgtgagct	aataaatacg	ggtggcctct
68941	aaaggcaaga	aagggtaaag	aactggattc	tcactctaga	gtcaccggga	aggaactatc
69001	aacatcttga	tttcagccca	gtgagactct	gtcagacttc	taagctacag	aactgtaaga
69061	taaatttgtg	ttgtttttaca	tcattaaatg	tgcagtaact	tgttacagca	gcaattagaa
69121	atgaatacag	aggactgggc	attaggcctg	tatctcagct	ttctctgatc	tcctgggtgt
69181	ttcctgttat	ttattgttgg	tttccccag	aatgagtgat	ctaagaggaa	gcaaaataga
69241	agccgcaatc	tctttatgac	ttagcctcag	aagacacaca	ctggggccag	gtgcagtggc
69301	ttatgcctat	aatcccagca	ctttgggagg	ctgaggcagg	aggaccactt	gagcccagga
69361	gtttgagaca	ccctggacaa	cacagggaga	ccctcactct	ataaaaaata	acaaaaatta
69421	gccagggtgtg	gtggtgcaca	cctgttagtc	cagaactttg	ggaggctgag	acaagacaat
69481	gacttgagcc	caggagtgtg	agacaggtct	ggacaacgtg	gtaagactct	gtctttataa
69541	acatttttaa	aattaggcgg	ggcatggtgg	ctcatgcctg	taatcccagc	aatttgggag
69601	gctgaggtgg	gtggatcacc	tgaggtcagg	agtttaagac	aagcttggcc	aacatggcga
69661	acacYcgtct	ctactaaaaa	Yacaaaaatt	agccgggcat	ggtggtgggt	gcctgtaatc
69721	ccagctactc	aggaggctga	ggcaggagaa	tcatttgaac	ccgggaggtg	gaagtgttag
69781	tgagccgaga	ttgccttctc	cactccaaga	gttataaaa	attttgacca	tattttcttc
69841	tagcattttaa	tcaattaatt	aattaatgtg	agacagtccc	actctgctgc	ccaggctgaa
69901	atgcagtgg	gcaatctcgg	ctcactgcaa	cctctgcctc	ccggattcaa	gtgattctcc
69961	tgctttagcc	tcttgagtgc	ctgtgattac	aggcaccagc	cactatgcgt	ggctgatttt
70021	tggtttttta	gtagagacgg	ggtttcacca	tattggccag	gctggtctca	aactcctgac
70081	ctcatgatcc	Rccctccttg	gcttcctaaa	gtgctgggat	tacaggcgtg	agccactgtg
70141	cctggccttt	ttctttcttt	tttttttttt	ttcattagag	atgagttgtt	gttatgttgc
70201	ctctaactcc	tgggctcaag	cagttctccc	accctggctt	cccaaagtgc	tgctgggatt
70261	acaggagtga	gccactgccc	ccagcctctg	acagtttttg	tgactagga	attttgggaag
70321	acaattttac	ctggctattt	ctggctcata	caattgcagt	cagatggtgg	ctaaagctgg
70381	aacaataagc	agctaaaaca	gctgaaagat	aacctagcat	tctctctccc	tttctctgag
70441	tagtctccga	acctatctat	gttgctctct	gcattgggcta	gcttgggctt	cctcacagca
70501	tggcagcctt	aaggctttta	tagtcagctt	ccaaaatggt	cctcagtgat	tcctgcttcc
70561	tggatgtgat	accattgtga	agctctctct	cacattgaaa	ggggctgaac	tggccatttg
70621	ggataatgca	gaaatgacag	tggtgtgact	tagaggctaa	atcatgaaga	tatttgggct
70681	tccatcttgc	tcctttgtag	atcactcatt	ctagacaaa	ccagctacca	tgatatgaaa
70741	gcactcaagt	aaccctaggg	agagaggtct	ccttagtgag	gaactgaggc	cctgtaagaa
70801	acgtgggtgt	gctgcagtca	agtgggcata	ggccaaagta	aacatccaga	gtgactcagt

70861	gagtttagag	tgcaggcata	tagctccact	tgttatcaca	gccgtgtagc	cataacatgg
70921	gaaggctcat	cacttggctc	tgagccactg	ttgtctgtaa	aaggataaat	tgccctgctg
70981	acactgtgca	cagggctcgg	cccaacatgg	cttgacatgg	gacatggctc	ttgtgcaggt
71041	gcttgtagcc	agagaaagag	agaaaaccag	agctgtccat	ctcggggaag	ccaagacaca
71101	gctcagctag	ctcatgcccc	gagggagaaa	gagtaaggct	gtgggggtgtg	gtggctcatg
71161	cccataatcc	cagcactttg	ggaggccaag	gcagggtgat	cacaaggtta	ggagtttgag
71221	accagcctgg	tcaacatggt	gaaacccctt	ctcaactaaa	aatacaaaaa	ttagctggcc
71281	atgggtggtgc	atgcctgtaa	tcccagctac	tcaggaggct	gaggcaggag	aattgcttga
71341	accaggagg	cagaggttgc	agttagccga	gatcacacca	ctgcactcca	gcctgggcaa
71401	cagcgcgaga	ctccatctca	ggaaaaaaag	aaaaaagaaa	agaaaagaaa	gagtgaagct
71461	gctgaccctg	aaggagagac	tgccacaca	gctgtgtgtg	tgtagggagct	gccggagtaa
71521	gcagctgaga	cagagcagac	agtgcgagag	taagatgttg	atgatgagag	agctgctgaa
71581	taaagccatg	tctcatttac	ctgctgtctc	tcgagtgttc	ttctagctcc	ctgcctcacg
71641	tccactgctt	cctctcacac	ctcagctggg	gctggacccc	aaccctgagc	atgacgggcc
71701	ttctgtcaac	aaccagcagt	aacctgctgg	gcatgtgagg	gagctacctt	ggaatcagat
71761	tctgtaaaac	agtcacgcct	tcagatgacg	gtagcattgg	ccaacatttt	gactgcactt
71821	catgagagac	cctgagccag	aaccccttag	attcctaacc	caaggaaact	gtgtgtgata
71881	agtgtttatt	gttttttttt	tttttttttt	tttttgagaa	agagtctcgc	tctattgccc
71941	aggctggagc	acagtggcac	aatcttggct	cactgcaagc	tccgcctctc	aggttcacac
72001	cattctcctg	aatcagcctc	ctgagttagt	gggactacag	gcacccacca	ccacgcctag
72061	ttatcttttt	tgtattttca	gtatgagacg	ggtttcacgc	tgtagccag	gatggctcta
72121	atctcctgac	ctcgtgatcc	gcccgccctc	gcctcccaaa	atgctgggat	tacaggcgtg
72181	agtcaccaca	cccggccagt	gtttattgtt	ttaagatatt	ggctaggcgc	agtggttcac
72241	acctgtaatc	ccagcacttt	ggaaggccga	agtgggagga	tcacttgagc	tcaggagttc
72301	aaagtcaaga	acagcctggg	caacatagtg	agacctgttc	tctatttaaa	aaaatgtttt
72361	tagatgttta	tgtttgagct	gggtatgggt	tggtcacgcg	ctgtaatccc	agcagtttgg
72421	gaggctgagg	tgggtggatc	acctgtgggt	aggagatgga	gaccagcctg	gccaacatag
72481	tgaacccccg	tctctactaa	aaatacaaaa	aattagctgg	gcatgggtgt	gggcgcctgt
72541	aatcccagct	actagggagg	ctgaggcagg	agaatcgctt	gaacccggga	ggcagaggtt
72601	gcagttagcc	aagatcgtgc	cattgcactc	cagcatgggt	ctatgttttg	gaggtaattt
72661	gttacacagc	aataaataat	tcgttagagg	caccagcctg	gccaacatgg	agaaacctt
72721	ctctagtaaa	aattatccgg	gtatgggtgt	gcatgcctgt	aatcccagct	acttgggagg
72781	ctgaagcagg	agaatccctt	gaacctggga	ggtggaggct	gcagttagcc	aagatcgcac
72841	cactgcactc	cagcctgggt	cacagagcaa	gactctgtct	caatttaaaa	ataaaataat
72901	aataatatag	ggcagtcaga	ctgcccacct	ggcagctcag	gactagcaca	tgtgtctcag
72961	aaagccagg	ggaagctaca	tattttatga	tctaaactca	gaagtcatat	agcatctgtt
73021	ccactgtaat	cacaagcctt	cccagttcca	aggggagggg	acatagactc	cctcacctct
73081	tgatacaaga	agtgtcaaa	ttatatggta	agaagttggc	caggccctgc	ttgtctctgt
73141	tgttcatgcc	tgtaatccca	gcactttggg	aagacgaggc	agatggatca	cctgaggtca
73201	ggagtttgaa	gccagcctgg	ccaacatggt	gaaaccctat	ctctacaaaa	atacaaaaat
73261	tgttggtgca	tgggtgtatg	cactgtgaat	cccagctact	tgggaggcca	aggcagaga
73321	attgcttgaa	gctgggaggc	agaagttgca	gtgagccgag	attgtgccac	tacactctgg
73381	cctgggttac	agtgcaagac	tctgtctcaa	aaaaaaaaaa	aaaaaaaaag	agagaagttg
73441	gttgggcccc	gtggctcacg	cctgtaatcc	caacactttg	ggaagctgag	atgggaggat
73501	gcgttcaggc	cagaagatcc	atcgtttaca	gcctgagcaa	cacaaggaga	tcccgctcct
73561	acaaaatttt	tttaaaaatc	agctgggtgt	ggtggcaggc	acctgtggtc	acagctactc
73621	gggatgctga	ggtaggagga	tcgcttgagt	cagggaggtt	gtggctgccg	taagccatga
73681	acatgccatt	gcattctagc	ctgggttaaca	gagttagaca	ctgtttcaga	aaaaataata
73741	aaataaaaata	aataatgttg	taggacaggc	gtggggctca	cgcttggaat	ttcagtgcct
73801	tgggagctag	aggcaggagg	attgcttgag	aacaggagtt	cgaggctgca	gtgagctgtg
73861	atcgaccac	tgcactccag	ccttgggtgac	atgagcgata	tcttgtctca	ataaaataat
73921	acatacagtt	ctctttttaca	tcgagtata	gtaaaattttt	aaaaatacat	tgaagcgct
73981	tagaaagccg	cctgactctc	cctctccctc	tccctctccc	tctccctctc	cgtctccgtc
74041	tccgtctccg	tctccgtctc	cgtctccgtc	tccctccacg	gtctccctcc	acggtctccc
74101	tctgatgccg	agccaaggct	ggacgggtgt	gctgccatct	cggctcactg	cagctccctc
74161	gcctgattct	ctgcctcag	cctgtcgagt	gcctgcgatt	gcaggcgac	gccgccagc
74221	ctcactgggt	ttcgtttttt	tttttgggtg	agacggggtt	ttgctgtgtt	ggccgggctg
74281	gtctccagct	cctagcccg	agtgtaccgc	cagcctcggc	ctcccggggt	gccgggattg
74341	cggacggagt	ctcgttcact	cagtgtctct	tgggtcccag	gctggagtgc	agtggcgtga
74401	tctcggctcg	ctacagcctc	cacctcccag	ccgcctgcct	tggcccccca	aagtgcggag
74461	attgcagcct	ctgccagcc	gccaccccg	ctgggaagtg	aggagcgtct	ctgcttggcc
74521	acccatcgtc	tgggatatga	ggagcctctc	tgcctggctg	cccagctctg	aaagtgagga
74581	gcgtctctgc	ccggccgcca	tcccatctag	gaagcgagga	gcgcctcttc	cccgccRcct
74641	tcccatctag	gaagttagga	gcgtctctgc	ccggccgccc	atcgtctgag	atgtggggag

74701	cacctctgccc	ccgcccgcct	gtctgggatg	tgaggagcgc	ctctgctggc	cgcaaccctR
74761	tctgggaggt	gaggagcgtc	tctgcccgcc	cgccccgtct	gagaagttag	gaaaccctct
74821	gcctggcaac	cgccccgtct	gagaagttag	gagccccctc	gtccggcagc	cacccccgtct
74881	gggaagttag	gagcgtctcc	gccccgcagc	cacccccgtcc	gggagggagg	tgggggggggt
74941	cagccccccg	cccggccagc	cgccccatcc	gggaggttag	gggctcctct	gccccggccgc
75001	ccctactggg	aagttagggg	ccctcttgcc	tgccagtcgc	ccccgtccag	gagggaggtg
75061	gggggggtcag	ccccccgccc	ggccagccgc	ccagtccggg	aggttagggg	cgctctgccc
75121	cgcccgcccc	tactgggaag	tgaggagccc	ctctgcccg	ccagccgccc	cgccggggag
75181	gggggagggg	gggtcagccc	cctgcccgcc	cagccgcccc	gtccgggagg	gaggtggttg
75241	gggtcagccc	cccgcggcgc	cagccgcccc	gtccgggagg	tgaggggtgc	ctctgccccg
75301	ccgcccctac	tgggaagtga	ggagcccttc	tgccccgcca	gcccggccgt	ccgggaggtg
75361	gggtgggggg	tcagcccccc	gccccgcggg	cgccccgtc	gggaggtga	ggggcgccct
75421	tgccccgccc	cccctactgg	gaagttagga	ccctctgccc	cagccagccg	ccccgtccgg
75481	gagggaggtg	gggggggtcag	ccccccgccc	ggccagccgc	ccagtccggg	agggaggttg
75541	ggggatcagc	ccccgcggcg	gcccagccgc	cagtccggga	gggaggtggg	gggatcagcc
75601	ccccgcctgg	ccagccgccc	cgccgggag	gtgagggcg	cctctgcccc	gcccggccct
75661	ctgggaagtg	aggagccccc	ctgccccggc	agccgccccg	tccgggaggg	aggtgggggg
75721	gtcagccccc	cgcccgccca	gcccggccgt	ccgggagggg	agtggggggg	gtcagccccc
75781	cgcccgacca	gcccggccgt	ccgggagggg	ggtggggggg	tcagcccccc	gcctggccag
75841	ccgcccgcgc	cgggaggtga	ggggcgccct	tgccccgccc	cccctactgg	gaagttagga
75901	gccccctctg	ctctgttgaa	ggcagccctc	tcgttaagag	tcctaccacc	tcctaatctc
75961	taagtaccca	gggacacaaa	cactgcggga	ggccgcaggg	tcctctgcct	aggaaaacca
76021	gagacctttg	ttcacttggt	tatctgtgta	ccttccctcc	actattgtcc	tatgaccttg
76081	ccaaatcccc	ctctgcgaga	aacaccacaag	aatgatcaat	aaaaataaaa	aataaaaaaa
76141	aaaaataaaa	aaaataaaaa	aaaaaaaaaa	gaaagccgcc	tgacctgtat	acagtattct
76201	gaaaaggggg	tcgcgaggtg	catgtccaac	ctccgcggcc	gggggcagca	gcgaggtccag
76261	gcccagccgg	ggcctagcga	gcgggtcaa	atgggggtgag	gcctgtgcca	gacctctcca
76321	cctcggtggc	agccgcagcc	tcctccgcct	gcggctcctg	tcacgcggc	ggccacgtga
76381	gcgccagatt	ctggcgacac	gacctgtgcc	agtcctttgc	tgctttgcgc	agcctgtcct
76441	ccccgccagg	agcacccctc	ccgctccctt	ttaccacggg	ctccagccgt	ggctgccttg
76501	gggctccgcg	cgccctggctg	taYtccagta	cggtgggaaa	gaacgggtg	gaatgggtg
76561	ggtgggggtc	aaagaggaaa	cccagagatg	cagggcgccc	ccttcccggtg	gtctgcccc
76621	aattgctcag	gcaggccagt	cacgtgaggg	cgctccctcc	ccaagtttat	atttattatt
76681	atttattatt	tatttttttc	accttcaagt	ttattattta	ttattttatt	atttattttt
76741	gagcggagc	ctctctctgt	cgcccaggct	ggagtgcatt	ggcacgatct	tggtcactg
76801	caacctccgc	ctcccggtt	caagcgattc	ttctgcctca	gcctcccag	tagcagggt
76861	tacaggtgca	tgccaccaca	tcgggcta	ttttgtatt	ttagttaa	cagagtttca
76921	ccatattggc	caggctgggtc	tcgaactcct	gacctcaggt	catctgccc	ccttggccct
76981	ccaaagtgtc	gggattacaa	gcattagcca	ctgcacctgg	ctaacctcca	agtttaaaga
77041	cagccgcccag	gcccagtggtc	tcactcctgt	aaccccaaca	actcaggagg	ctgagggcag
77101	gaatttgaga	ccagcctggg	caacatagcg	agacccggcg	tctaagaaaa	atagggcagg
77161	cacggtgggt	tacgtctgta	atcccagcac	tttgggaggc	tgtggcaaaa	ggattgcctg
77221	agggggaaaa	aatcacccctg	ggggtagtg	tgcacactta	cagtctcggc	tacttgagag
77281	gctgaggtgg	gaggatcatt	taagtccggg	gctgcagtga	gctactatgg	agcgactgta
77341	ttacagcttg	agcaacagag	cgagacccca	tctccaaata	aataaataaa	gatagcctcc
77401	aaagatgtca	cttgcttcac	ttagcacttt	ttattgaaca	tatttaggaa	atatataacc
77461	atggtaaaga	aaattgcaat	tagtacaat	acacactaca	cacatatgca	cgtgcacaca
77521	ctgaaacgtg	tccttccag	cgctcctgc	ctggataatt	ttttttttt	tgcaacgga
77581	gttttgctct	tggtgccag	gatggagagc	agtggcgggg	tatcggtcca	ctgcaacctc
77641	ctcctcccg	gttcgagcga	ttctcctgcc	tcagcctcgc	gagtagctgg	gattacaggc
77701	gccagccaca	acaccgggt	gatttttgta	tttttagtag	agacgggggt	ttgccaagtt
77761	ggccaggctg	gtctggaact	cctgagatcc	gcccacctcg	tcctctcaaa	gtgctgggat
77821	tacaggcctg	agccactgcg	cctggcctaa	agtaattgtc	ttcttattgg	ttctctgccc
77881	tttgggtctca	ccaaacgccc	cactctaaat	cactgcagac	aaggggttct	gtctaaggag
77941	gagagcccac	cagttaaaac	ccttcagtag	ttcctaacta	ccctaggaca	aatgcagact
78001	tatccttcgc	tcctgctgc	cgccctctc	cttccagat	cccatatggc	tcacgcata
78061	tcctcagacc	atctgggctg	ttgtctgtcc	tgccacacac	ccttcccacc	ccgtccctt
78121	gcaagtccca	ctcaggctgc	acctacacag	ctgtctcagt	ttctatctga	ggctcccgc
78181	gcctcctctg	aaagtccca	tcacagccaa	tgatactgaa	actgttttct	tatgcagggc
78241	acggaaagat	ctattcaact	cactgcccac	tctcctctcc	tggtgcagaa	aagacaccca
78301	gtcagcgact	tgcatattct	gcaggcatga	tgaaataata	ataatgccta	acccttatat
78361	agtgtcaatt	ccacgcctgg	cactgttcta	ggcactcata	taaattcatg	taatccacac
78421	aaccccaaaa	ctgatgatat	ctctctcatc	taccaagcac	tgagcagtta	aataacttgc
78481	tccagatatt	aaggggtcga	gctgggggtt	gaagcctggt	taccataaaa	tgaaccaaga

78541	actggaagga	ggacaagaSc	tccgagaagg	agtcaggtag	ggcgtgatct	gtgcgcttta
78601	catctaagat	cttccagctc	ccagggagcc	cgtttcatag	agcaggagat	agaggctggg
78661	agggacacgg	agagcctcga	gagccgtgtg	gaggaagcgg	tgctgttttg	ggtccgggag
78721	caagggcgtg	gcctggatgc	gcgggcgcgc	gggacggcac	gtcctcagac	caaactacaa
78781	ctcccaggac	ccagcgggcg	ctgccgccca	cgcgacgtca	cggcggcgga	gggcgcaggc
78841	ggctgggdcg	ctggcgagtg	gactgttcga	gcccttcgcg	tgggaccggg	gcccttggctc
78901	cggccccgcg	gtaagtgggg	cgaccccagc	ctactcagtc	cgcggaggcc	ccgcggcgca
78961	cgtccgcagc	ctccatcaca	gcgcggggcg	gcagacgggg	ctggcatcta	ccatatgggg
79021	ggcatccggg	ccgaaccaag	tgaccgcggt	ggggggtccc	gctggggact	ccgtgccgca
79081	ccctcccaag	ccggccccag	gggcccaggg	ctggtgtcgc	acgttcgctg	gccgcgctcc
79141	cagggccccg	gtttgaaggc	gctgggcagg	caggggcagc	cccggcccct	gagaagggta
79201	cccgggaccc	cggggcgctg	ggcgcgagtt	ttcgggctgg	aagggtctga	ggggctcctc
79261	ccccgacagc	cctcccaccg	ccagtagagc	ctcgggttgg	ggaatagaag	cccccgggag
79321	gctaggtcct	ttgggcgcgg	cctgtgtgca	tctggggaga	cgggtggagt	ggtggggaga
79381	ggtcgcccg	gtctggggag	accgatgcac	aggtggagag	atggtgcggg	ttctgtggat
79441	tcggatcctt	acaacttctt	cttccccgcg	ccggtagatg	ggagctgctc	tccgcgggct
79501	gagcctgtca	gcatectcga	cgcaccctgg	tccttgaagt	cggagaagMg	ccccaccaca
79561	cccacacccc	cttgccccat	tttgggtcgc	ctgggtcctc	agtcctagcg	gatacctcagt
79621	cctagcggcc	accgggtctg	aaaggagcaa	gacgatgatc	ctggcgctcg	tgctgaggag
79681	cgggtcccggg	ggcgggcttc	cgtcccgccc	cctcctggga	cccgcactcg	cgtccggggc
79741	cgctctcgag	tcggcccaccg	acacacacca	cgtggagatg	gctcgggagc	gctccaagac
79801	cgtcacctcc	ttttacaacc	agtccggccat	cgacgcggca	gcggagaaag	tgcgcaaggg
79861	ggcagccagc	ccagggtccg	gRatgtaggc	gggagggaga	gtgttggggg	ttctctgctc
79921	aaggcctctc	tcctctctta	gccctcagtc	cgcctaacgc	ccaccatgat	gctctacgct
79981	ggcgcgtctc	aggacggcag	ccaccttctg	gtaagattca	cgcctcttat	tttctctcgtg
80041	gatcttggag	ctctcccaga	cactcaggct	ccagcccgcg	cttcccttct	cattttctcc
80101	cagaaaagtg	ctcggtacct	gcagcaagaa	cttccagtga	ggattgctca	ccgcatacaag
80161	ggcttccgct	gccttctctt	catcattggc	tgcaacccca	ccatactgca	cgttggttaag
80221	tagagaggac	cttaggtcag	cgggccaccc	tgccccgggg	gcaagtgggg	agtctggggc
80281	ccagagtggc	agacgattgc	ttgcctaaag	gtgtcagggc	cacacaggat	tcaacccag
80341	gccttcagaa	gccaaagggtg	tgtattcaag	gagcctggaa	gggtcgaagt	gggggtttga
80401	tcacgtggtc	gcacagctgg	gtggtgatcc	ccatgggtag	gtgggggtgg	ctgttctctg
80461	ctcagtgcgc	atgcggcttt	gtgaattccc	acacctcttc	cttgagcat	gagctatata
80521	tcogtgcctt	ccagaagctg	acagacttcc	ctccggtgag	tgctggggca	gagcaggggtg
80581	aggggctgag	aggttgggct	tggaccaccc	ttcctcatga	ctctgtgacc	tgcatatcaa
80641	ggaccaggcg	gacgaggccc	agtactgcca	gctggtgcga	cagctgctgg	atgaccacaa
80701	ggatgtgggtg	accctcttgg	cagagggcct	acgtgagagc	cggaaagcaca	tagaggttgg
80761	ggcagcaaag	gagaggccgg	gcctgctggg	ggtgggaagg	gcacgggatt	ctgagacctc
80821	actctttaca	ggatgaaaag	ctcgtccgct	acttcttggg	caagacgctg	acttcgaggc
80881	ttggaatccg	catgttggcc	acgcataccc	tggcgctgca	tgaggacaag	gtggggctct
80941	gggacctgag	accacactgg	gaacattaaag	tgagacagag	gagactgggc	tggggatccg
81001	ggtcaagggc	ctgggggctg	aggctgtggg	gctggtgctt	tggggcagtt	ccgaagtggc
81061	cagcatcttg	gggtggggct	aggggcgtgg	gtagtcttga	cctcctttct	ccggccagcc
81121	tgactttgtc	ggcatcatct	gtactcgtct	ctcaccaaag	aagattattg	agaagtgggt
81181	ggactttgcc	aggtgaggca	agaatggctc	agggggtggg	cagacatctg	gggcagggaa
81241	ggcttgggtc	tgagcccttg	cccggggcat	gatctgcggg	gagcaggggt	tctcaacccat
81301	ggcactattg	acattttccag	ccagataatt	ctttgtcaYa	ggggctgccc	cgtgcacgctt
81361	aggaagttca	gcagcatccc	tggcgccagc	agtactgcct	agttgtgaca	aacaaaaatg
81421	tctctgcaca	ttgccatatg	ttacttaggg	gggcagaatt	gtttccagtt	gcaaaccact
81481	ggtggagggg	cccctgactg	aaccctcgct	cctatccgca	gacgcctgtg	tgagcacaag
81541	tatggcaatg	cgcgccgtgt	ccgcatcaat	ggccatgtgg	ctgcccgggt	ccccttcac
81601	cctatgccac	tggactacat	cctgccggag	ctgctcaaga	atgccatgag	gtgggggtggc
81661	ttgatgtgct	ggcttggggg	Yggacaggaa	ccgggNtgct	tgtacctact	ggtctttccc
81721	ctctgcatag	agccacaatg	gagagtcacc	tagacactcc	ctacaatgtc	ccagatgtgg
81781	tcatcaccat	cgccaacaat	gatgtcgatc	tgatcatcag	gtttgccttg	agtgggagtt
81841	gagctgaggt	ggatgggatg	gggtcttagg	cactgtttct	gacttgattt	aggaccttga
81901	gcccccttct	gccccattct	gggacttggg	ccctgaccag	acaaactatt	ctctgaatcc
81961	tgagatggcc	atgagctgct	tattaatgga	tctggggcca	gctgcaggcc	taggtatcct
82021	gcctctgtca	gcagctgagg	agcttgaaat	tgagaaatag	tcaggagtcg	gtctaggatg
82081	ctgggcccag	gataaatgtc	acatcctgtg	agaaggtata	agcagtcagt	ggccctggca
82141	tggggtgagg	tgatataaac	aaggcccag	ggtctagggtg	gaccacattc	cagctctggg
82201	tgggaaggaa	aggaaggcag	actttgcact	gtctgcttgg	gggggtgggtg	gtaccccatc
82261	aaagctgagc	caagcccatt	gttgttgcca	tcttgctagg	atctcagacc	gtgggtggagg
82321	aatcgctcac	aaagatctgg	accgggtcat	ggactaccac	ttcactactg	ctgaggccag

82381	cacacaggac	ccccggatca	gccccctctt	tggccatctg	gacatgcata	gtggcgccca
82441	gtcaggaccc	atgcacgggt	gagaccctgc	caggccagga	tggaggggtg	ggggacccca
82501	ggagactcaa	gcctctgaag	cctcctgtcc	tgtccccctg	cccaccccca	gctttggctt
82561	cgggttgccc	acgtcacggg	cctacgcgga	gtaccctcgt	gggtctctgc	agctgcagtc
82621	cctgcagggc	attggcacgg	acgtctacct	gcggctccgc	cacatcgatg	gccgggagga
82681	aagcttccgg	atctgacccc	acagcctttg	gcctgctcac	ccgaccagcc	tgggccgcat
82741	tccctgcagg	acctcccggg	tcaggcaggg	cggccccctg	ctccacacac	tgtctgatct
82801	tgggtctcag	ggacccagac	agatggactt	acatggagct	gggcaactgc	ctgcctcaac
82861	agggtccatt	gcctcctcgc	ctccagaact	tggagcaggg	aagtgggcac	cctgaggcct
82921	ccagcaccag	ttccgtcatt	ctcgttctct	gggaaccccc	actctgacct	gttattaaag
82981	ttcacatttt	gaatgccctc	tcgggccccg	tgtgtgggga	gggcagggtg	acttttgttt
83041	ctgcccccat	tcaggttcac	tgagcccttg	ggttgaactg	gttcgtgtcc	cagtctctta
83101	cctgccctga	gagcctggca	ggccaggagt	agaatgggtc	ccaagtctgt	tgcattgttt
83161	atttggtggg	agtgggatga	ctgcagcacc	ttatacaaa	agctttcatt	catcttgttg
83221	aacaaatggt	tccgggtccc	agataatatt	gaaggccag	actgaccag	cttcgggcat
83281	cagttttgac	tcttcctttc	ctggcagtc	cagtttctag	aggtgaaggt	caccagactg
83341	ggcaaactcc	tgagccaact	gcttcccaag	cctgagtagg	ttaaaaatac	tgtgtctgct
83401	gctgccaagg	aaaagaacat	acaaggttgt	gccttggcag	gccctagcag	ggactgggtg
83461	ccccactgca	aggaaagggt	gggccctgat	agaaaggacc	aaggatttgg	gcaaagRtat
83521	caggttaggt	caaggttaga	cctgaatcag	aactccagat	gacatcttag	gtaggaacac
83581	cctaccaccc	ttgccaggga	agaaaggcct	aaggcgccgc	tgggtggggt	gggaggagaa
83641	ctggaaagtt	ctcttgccct	cacatgtgag	ctccacagc	aaacttccct	aggctggctc
83701	taggcctgta	ccatctccta	cccttcacgg	ggatggaggg	gaagtgttat	gtggaagcca
83761	aatggcaggg	gctaggaaac	cacagtgact	tgctagactg	aaaaatcccg	ccagctgcaa
83821	ggcagggtgc	tgaggctgga	gaggcaggga	gcagtcagag	gccagggccc	tgaaacatgg
83881	gagttatctt	gagccatagg	gatccatggg	tgagttttta	tttattttag	aatgggttct
83941	tgtctgtgtg	cccaggctgg	aatatggtgg	ctgcagagtt	caactgcagcc	ttgaactcct
84001	gggatcaaga	gattctccca	cctcagcctt	ctgagtagct	tggaccatca	tgccaggcta
84061	aatttttaaaa	ttttttgtag	aaacagggtt	tctacaaagc	cctatgttgc	cccgggctgg
84121	acttgaactt	ctgggctcaa	atgatccttc	caccccagcc	tcccaaagtg	gtgggggttac
84181	aggcatgagc	cactgcagct	ggcccatgag	tgggttttga	gctgggaagg	gctgggttctg
84241	gttgaggatcc	ctgagaggat	tcattgtccac	gtgatttctt	aagaaagtgc	tcccagaaca
84301	gagtagggga	agtaggaagg	ggaaggggag	gaagccaagc	aaggatgtga	cctcaggcaa
84361	aagcccagaa	ccagtcaatt	atgcctcagg	gttgaaaggta	agagagctaa	acctcagagt
84421	tactgattaa	tttctccact	tggcagtcac	tgggttaaagt	cagttgggaa	agtgaacagc
84481	tctatttaacc	taaggatggg	tttttaagaa	gagcctcagg	tgctggtgtg	ggtccttgaa
84541	agcacatcaa	aggtaatctg	ggcacacaga	aacagcaaga	actcccagag	gatctgggtg
84601	gagcacctac	attgtttttt	tgtttgtttt	gtttcgtttt	gtttttttta	acggagtctc
84661	aatctgtttg	tcaggctgga	gtgcagtggc	tggatcttcg	ctcactgcaa	cctccgcccc
84721	atcccccccc	aaccccaggt	tcaagcgatt	ctcctgcctc	agcctcccga	gtagctggga
84781	ttacaggcgc	gtgccaccac	accagactaa	tttttctatt	tttagtagag	atggggtttc
84841	accacgttgg	ccaggctggg	ctcgaactcc	caacctcgtg	atccatccac	ctcagcctcc
84901	caaagtgcc	ggattacagg	catgagccac	catgcctgtc	ggatgtttct	tgatttgtaa
84961	cctctgagag	acccatccgc	aggccctgag	cattccactc	ctctcagaat	tgtttccaag
85021	cccaataacc	acattataaa	tcaaacaaga	ttcagagaat	agccaaaggg	aatgtttact
85081	gagtaacctt	ccggtctggc	actttgcaat	acacttgat	attgctaaga	cggatgttcc
85141	aaccgttaca	tagttatatg	attgatagtt	atacatgctt	aactgctggg	gattgggttc
85201	aggaccgcct	gtgaataccg	aaatctgcag	gcgctcaagt	cctacagtgt	gccctgccaa
85261	acagcagata	tgaagtacgc	tcttcagatc	tgtgggttct	gcacctttac	aatatttctt
85321	ttcctttctc	tttcttttcc	tccttctctc	cctctttttt	ctttttcttt	tttgagatgg
85381	agtctgtgtg	tgtcgccag	gttggaagtgc	agtggcgcg	tctcggtcca	Mtgcaacctc
85441	cacctcctgg	gttcaagcag	ttctcctgcc	tcagcctccc	aagttagctg	gattacaggc
85501	acacgccacc	acccctgact	gttttgattt	ttcagtagag	acggggtttc	acaatgtggg
85561	ccaagctggt	tttgaactcc	tgacctcaag	taattccact	gcttcggcct	cccaaagtgc
85621	tgggattaca	gggtgtgagc	accgcgcccc	gtcttttttt	tttttttttt	gaggcagagt
85681	ttcactcttg	ttgccaggc	tggagtgc	tgccacaatc	tcagctcacc	acaacctctg
85741	cctcccagg	tcaagcgggt	ctcctgcctc	agcctcccga	gtagctggga	ttacaggcat
85801	gcggccacca	cgccctggcta	attttgattt	tttagtagag	atggggtttc	tccatgttgg
85861	tcaggctggt	ctcgaactcc	cgacctcagg	tgatctgcct	gcctcggcct	cccaaagtgg
85921	tgggattaca	ggagttagcc	actgcgcccc	gcctcctttt	ctttcccccc	tttttctttg
85981	agacagggtg	tctgtcacc	aagctggagt	gcagtggagg	gattatagct	cactcagcct
86041	cgacctctg	ggtttaagcg	atccctctgc	ctcagcctcc	tgagttaggt	ggactacagg
86101	tgcgggcccc	gaggccccag	taattttttt	tttcccccaa	attttttagta	gaaaggaggt
86161	ctctatgctg	cccaggctgg	tcttgaactc	ctggcctgaa	gcgatcctcc	tgcttgattt

86221	cctgaagtgc	gagattacag	gtgtgagcca	ccatacctca	acactgtatt	ttcaacccgc
86281	tcttcggtca	atctccaaaag	gtgggacatg	cggatatgga	gggcccattg	tRtatgggtg
86341	gaccatacac	atataaatgg	ctttaacctt	tactgactct	cacagaaccc	tcagtgcagt
86401	ggcgtgatct	cagctcactg	caagctccac	ctcccgggtt	cacaccattc	tcctgcctca
86461	gcctcccag	tagctgggac	tacaggggccc	cgccaccacg	cccggctaatt	tgttttgtat
86521	tttttttttag	tagagacgga	gtttcatcgt	gttagccaga	atagtctcga	tcttctgacc
86581	tcgtgatcca	cccgcctagg	cctccaaaag	tgctgggatt	acaggcgtga	accaccgcac
86641	ccggcctttt	tatttttttt	gagatggagt	ctggctcttg	gtcccaggc	tggagtgcac
86701	tggcgggatc	tcggctcact	gcaacctccg	cctcccgggt	tcaagcgatt	ctcctgcctc
86761	agcctcccga	gtagctggga	ctacaggtgc	gtgccaccac	gcccggctaa	attttgtatt
86821	tttagtagag	acggagtctc	acggtgttag	ccaggatggg	ctcgatctcc	gcccgcctcg
86881	gcctctcaaa	gtgctgagat	tacaggcgtg	agccaccacg	ccccgcccaa	ctcgtccttt
86941	cttttagactt	tatcctgtga	gggtgaatta	tggcctgtcc	ctggacacac	ccgttctgct
87001	ttccccgcac	caactgtatc	ccaaataggg	gaagttagct	cttcaacctt	caaaaatggg
87061	gcactggctg	ggcacgggtg	ctcacgcctg	taaccctagc	actttgggag	gcccaggcgg
87121	cgcgatcacc	tgaggtcagg	agttcgagac	cagcctggcc	aacagggtga	agccctctct
87181	ctaataaaaa	tacaaaaaat	agccgggcgt	gggtggcgcg	gattgtaatc	ccagctattc
87241	aggaggctga	ggcaggagaa	tcgcttgaac	ccgggaggcg	gaggctgcag	tgagccgaga
87301	tcgcgtcact	gcactccagc	ctgggcgaca	gagcgagact	ccatctcaaa	aaaaaaaRag
87361	gggggggggc	ggggagcgac	attaggccag	cgcgagtgtg	cggctccagg	ccaccagggc
87421	ggcgctctg	gtctctccct	gttaaagggtg	gggcagcgcc	cgctgaaaaat	gggcacatct
87481	ttgcttccgc	aggtaccagg	gcctttaaaa	aggttcagct	ggccgggcac	ggtggctcac
87541	gcctataatc	ccagcacttt	gggaggttga	ggctggcaga	tcacgaggtc	aagagatcaa
87601	gactatcctg	gacaacatgg	tgaacccccg	tctctactaa	aaatacaaaa	attagatggg
87661	cgtggtggcg	cgcgccctga	gtcccagcta	cccgggaggc	tgaggcagag	gaatcgcttg
87721	aaccYggagg	cggaggttgc	agtgaagcga	gatcgcgcc	ctgcaactca	gcctggcgga
87781	cagagcatga	ttccgtctca	aaataaataa	ataaataaat	aaataaataa	attaaaaaaa
87841	ggttgagttg	agtcacgcga	tgaggtaagt	ctgtctcttc	tcctcaccca	taggctgcag
87901	ccacaggtcc	gcacggaaaa	ccaggagctc	caagactgcc	gcacactcaa	catccaactc
87961	gcggggccgc	ggggagcgcc	acagaagcca	atcggtcctt	gctacatcac	agaccccgc
88021	ctcaaacctt	gaggagcccc	ccccagcgac	tgctgacccg	gagccaaaag	gagctcggtg
88081	cgtccgattc	tggtcgctact	tcccttcccg	cgatggcggc	acaggggagct	gctgcgggcg
88141	ttgcggcggg	gacttcaggg	gtcgcggggg	agggcgagcc	cgggcccggg	gagaatgcgg
88201	ccgctgaggg	gaccgcccc	tcccggggcc	gcgtctctcc	gcccagcccc	gcgcgcggcg
88261	agccgggaagt	cacgggtggag	atcgagagaa	cgtacctgtg	ccggcgaccg	gatagcacct
88321	ggcgtgaggg	cggggcccag	ggctgggggg	ggggcgagag	tcaggggccag	ggggggggcg
88381	ggggcctgag	gacaggctgt	cagtgaggcc	aagatccggg	ggctgggagt	ggagaggagg
88441	agggggcggg	cttgaggaaa	gaacgcgcg	ttccggcgcg	tgagaaacca	gcccgggtgt
88501	gggaggctgt	tgaccttgaa	ttatgccgag	cgacgcctac	aaacccaccg	ctcaggcctt
88561	caccaggatt	gttcccattc	cacttccctg	cccagttcta	ggcttccatt	ctttttcttg
88621	ctaacgctgc	ttctccaccc	tctcttgtct	ctgcgtcttc	tttttccatt	tgctcctggc
88681	agccatccgc	agagagaaga	ccttccagaa	acaacaggct	tccctttcct	aaagtctctg
88741	tgcttctctt	cattttcaaa	attaacccca	aactccttag	ggtggcattc	atttttgtga
88801	cctctccagt	ttctagccaa	cactaggaaa	gggcattgcc	aggccagaa	acactgtgcc
88861	ctctgaagac	cacacgcctt	ttaccacctg	tgccctttgc	ttggaatgct	ttttcttccc
88921	ttctccttg	tttgccctgc	tagctcctac	tcctcctctt	agcttcatat	cctttgtgat
88981	gtcatccttg	attcccttcc	aggcaaatg	agtggttccc	ttctctatgt	tcctgcaaca
89041	tttttttcc	aYctcagtc	tagtttttgt	aacattatgt	tgtaattttc	tctctctgtc
89101	ttcctccatc	atactgggaa	cttctggagg	gcagcacttc	ttgtgattca	tcactgtgtc
89161	ctctgtaccc	ggcaacaaca	cagcataggg	ccagacacgt	agtgggtgcc	ttactcattt
89221	attgaatgct	ctgctaattg	tcaggtgcgc	tgctaaaaat	tttgagggca	tttaattcac
89281	ttctaagggt	ggtactatta	tcccttttgt	cttttttttt	tttttttttt	ttgagacagg
89341	atctcgctct	gtcaccacag	tgggattgcc	gtggcacgat	gacagctcac	tgaagcctag
89401	accttctagg	ctcaagtgat	cctcccacct	cacctctca	gagtgttggg	attactggca
89461	tgagccactg	tgcccagctg	ttatcccttt	tcacagatga	agagactgag	gctcagaaa
89521	ttggaataac	ttgctcagtt	acacatagct	aggaagtagg	gagctggaat	tttgtgtatt
89581	tttttttgtt	tgtttgtttg	agatggaatc	gctctctgtt	gccaggctgg	agtgcagtgg
89641	cgccatcttg	gctcactgcg	atctctgcct	cccggattca	agcgattctc	ctgcctcaac
89701	ctcctgagta	gctgggacta	cagttgtgcg	ccaccacacc	cagctaattt	ttgtattttt
89761	agtagagac	aagtttcacc	atgttggcca	ggtatgggtt	gacctcttaa	cctcgtgatt
89821	gacccacctc	ggcctcccaa	agtgtgggga	ttacaggcgt	gagccacgat	gcccagcctt
89881	ggaattatag	attaatgtag	gttatcttaa	tccagagcca	gcactcatga	tccctgaaca
89941	aatgaacaac	tgccgggtatg	tgttaaaatg	agcattgaat	ttagagttag	acttaaatct
90001	agttaagtat	tagctcttaY	cacttttcaa	atctgtgacc	ctaggttaagc	ctcagttttg

90061	ttctgttttt	tggggttctt	ttttttgaga	cagagtcttg	ctcttgtcac	ccaggccaca
90121	gtgcagtggt	acaatctcgg	ctcactgcaa	cctctgcctc	ccgggttcaa	gcaattctcc
90181	tgccctcagcc	ttgcctccca	ggttcaagca	attttcctgc	ctcagcctcc	Ygagtagttg
90241	ggattacagg	tgtgtgccac	catgcctggc	tttttttttt	ttcRtctttt	tggtacagat
90301	ggggttttgc	catgttgacc	aggctgtgct	caaactcctg	gcctcaagtg	atctgcccac
90361	ctcggcctcc	agaagtactg	ggattacagg	tgtgagccac	acaccagcc	agttattttc
90421	tcattcgtaa	aatgagaata	acagccYcaa	cctgatccac	ctcacagttg	tggcagtgat
90481	tgagtagaat	ggcagaggtg	gaccagtgag	gttagctgat	gtgtggccct	gaactctgga
90541	aactgcttct	cagtgtgcag	gattcctttt	gttccagcct	taccttctctg	atgaccctag
90601	Nctttttcca	tcacagattc	tgtctgaagt	atccagtctc	gagtgaacga	ccaggagggc
90661	cgagaggaat	tctatgtaca	ctacgtgggc	tgtgagtggc	ttggcacatc	tgggcgggg
90721	tgCRgggagt	tggtctgccct	ggcagcactc	ttatggccca	tacttacagt	taaccggcgg
90781	ctggacgagt	gggtagacaa	gaaccggctg	gcgctgacca	agacagtga	ggatgctgta
90841	cagaagaact	cagagaagta	cctgagcgag	ctcgagagc	agcctgagcg	caagatcact
90901	cgcaaccaaa	agcgcaagca	tgatgagatc	aacctagtgc	agaaggtccg	gatcccttcc
90961	catccacggg	cccaggaggc	ccagcttctc	tgccagttcc	cttgggtctc	tcgggcccc
91021	gtgccaaaac	catagcaaat	cccatttctt	aagctcctgt	agtgtgtcag	ggactttacc
91081	tacttctctc	tacttacttt	tcataggtaa	gaaaacagac	agagggtaa	ggcttcgctg
91141	cagtctcaca	gcaaggaaat	ggcacagctg	ggactagaac	tcaggcctcc	ggcactggta
91201	gcctcttttg	tctcctagta	gctgtgtctt	agtagaaaa	atgacttgaa	gttcagtggg
91261	aaagacgaga	atgttattcg	gcagttgcta	tagccaagca	ccctgccaa	tgctgtatat
91321	gtataagctc	ctcaaatcct	tccaacaat	tcaacacaca	ggaatagtta	cttttttgg
91381	aggtgaggaa	gcaggctgag	agagattaag	taacatgcc	aaggtcacac	agccagtaaa
91441	acgggtgggt	ggaattagaa	taataatgcc	aaacacttgt	tagccggtga	gcagaagtat
91501	tcagctaaat	ggtttttgtg	tgtgtgtgtg	tgtgtgtgtg	tgtgtgtttt	gttttttgtt
91561	tttttttttg	agacggagtc	ttgtctgttc	gccaggttg	gagtacagt	gaatgtctc
91621	ggctcactgt	aaactccgcc	acctgggttc	aagcaatttt	cctgtctcag	cctcccaagt
91681	agctaggact	acaggtaact	gccaccacgc	ccggataatt	tttgtatttt	tagtagagac
91741	gaggtttcgc	catattgggt	aggctgggtc	cgaactcctg	acctcaggtg	atctgcccac
91801	ctcgccctcc	caaagtgtct	ggattacagg	cgtgagctac	tgcgcccgcc	ctaagctaa
91861	tgttttaata	catgacctca	ctgacatttc	acagtaggcc	tgtgaagtaa	tgatgcagaa
91921	gagctacaga	taagtttagg	gaccagaca	ggaatttgat	caaccagcta	aattgtgcta
91981	cctcagagcc	agaggcctta	gcaatgattt	tggttttttt	cttaagagac	gggagtcttg
92041	gtgtgttgct	caggctggac	ttgaactcct	gggctcaaat	gttcttctctg	ccttagcctt
92101	ctaagtagct	gggactacag	gagtgcattc	ctgtacctgc	cttgggtttt	tcttgtattg
92161	atacttaata	attgtacata	ttattgttat	tttttttttt	tgagacagcg	tcttactcta
92221	tcgtctaggc	tgagtgcagt	ggcacaatca	tggtcactt	cagcctcaac	ctcctgggct
92281	caagcaatcc	tctgcctca	gcctcctgag	tagctgggac	cataggcacg	tgccaccatg
92341	cccagcta	tttttatttt	ttgttagaga	cgggactctg	tatgttgccc	acactagtct
92401	caaactcctg	gtctcaagcg	atcctcctgc	cttagcctcc	caaagtgtctg	ggattacagg
92461	tgtgagccac	tgcgcccgac	caagtctcag	tttctaaatt	tataaaaaatc	aggagtgcac
92521	accagagcaa	gatttgtcaa	agtctccaga	caccagaya	tttgttactt	tcagctattg
92581	cagcttcatc	cttctgagta	catgagttag	tgggcaagcc	caacatttaa	atttctttta
92641	atggaagaaa	gaaagtggat	aagatagaaa	aatcagggtg	ttgcacttta	ttaagtgtag
92701	gtattgtttt	gagaagcttt	tatttcacac	ccaataaaag	tgtgcaggat	tctcttgtt
92761	ccagccttac	cttcccgatg	atcctagcac	atacttgtgt	gtatatacag	acatgttaca
92821	gtgccaaagt	tacttcctac	tgtgtgtgta	tttaaaaaat	tgaaataccc	taaaactggaa
92881	gaactcctca	aggtctaaac	tgccgaatcg	aacttcctat	tgtaatggaa	atgttctata
92941	gctgtactat	ctagtaccgt	agccactagc	cacatgtagc	tgttgagcac	ttgaaatgtg
93001	tctagtgtga	cgaagaatta	agaatttttag	ctgggcgcgg	tggtctacgc	ctgtaatccc
93061	agcacttttg	gaggccaagg	tgggtggatc	accaggtcag	aagatcgaga	ccatcccggc
93121	taacatgggt	aaaccccatc	tctactgaaa	aatacaaaat	aattagccgg	gcgtgggtggc
93181	gtgcacctgt	agtcccagct	actcgggagg	ctgaggtagg	agaatggcgt	gaacccggga
93241	ggtggagctt	gtagttagcc	gagatcctgg	gctcactcca	gcctgggtga	cagagtga
93301	ctccgtctca	aaaaaaaaaa	aatttttaaa	ttatgggtgc	agtggctcac	gcctgtaatc
93361	ccaacacttt	gggaggcgga	ggtgggtgga	tcacttaagg	tcaggggatt	gagaccagcc
93421	tgcccaacat	ggccaaaacc	ccatctctac	taaaaataca	aaaattagcc	gggctgtgtg
93481	gtgcacacct	gtaatcccgg	ctacttagga	ggctgaggca	ggagaatcgc	ttgagcccag
93541	gagatggagg	ttgcagtcag	ccaagatcat	gccactgcac	tccatccagc	ctgggtgtcc
93601	gatcaagact	gtctcaaaaa	aaaatttaatt	atcatttttt	aaaaaaaaatt	tatttttatt
93661	tttccagagc	agagtctcac	tctgttaacc	aggctggagt	gcagtgtgtc	gatctcagct
93721	cactgcaacc	tccgcctccc	agattcaagt	gagtttctctg	ccgcagcctc	ccaagtagct
93781	gggactacag	gcacgcacca	tcatgcctgg	ctaatttttg	tatttttagt	agagataggg
93841	tttcaactatg	tttggcctca	aactcctgac	ctcaactgat	ccgcctgcct	cggcctccca

93901	aagtgtctggg	attacaggtg	tgagccaccg	tgcccagcca	aaattttcat	tttttgagac
93961	agagtctcac	tctgtcacct	aggcgggagt	acagtggcga	gctcttgggt	tactgcaacc
94021	tccgccttcc	tggtttaagc	aattttttgtg	cctcagcctc	ccaagtatct	gagattacag
94081	gcgtgtgcc	ccacacccag	ctatttttta	ttttttatatt	ttagtagaga	tggggtttcg
94141	ccatgttggc	cagactggtc	tggaaactcct	ggcctcaagt	ggtcctccca	cctcggcctc
94201	tcaaagtgc	gggattacag	gcgtgagcca	ccacgttttag	ctgagtataa	ttttaaatag
94261	ccctatgtga	caagtggcta	ctttatttga	cagtgtagat	ctaagattaa	ttcctcaact
94321	gttttgcact	caacaaagac	atacctctga	gttggcaacc	agcaggggtg	ataacgggcc
94381	agtgggtgata	aaatcaaaga	ataggtaatg	aaacaatcat	ccagttaaca	atcagcaagg
94441	ttcttcagag	cctaattaat	gtttaattct	aaataaattg	caacaattaa	gaaaagtagt
94501	gtttttgtag	ttcttttattg	catttgcaga	agagtactgt	attttgtgag	aattgggaaa
94561	tgtgttttaa	agagatgtgt	gaggagga	atctttcttc	ctcatcttta	cagtgggttag
94621	gttcagaatt	ggacagacca	ggggcccatg	cccaactctg	ctgagtacca	gttctgtgac
94681	cttgggcaag	tgacctcacc	tctttgtgcc	tcagttttct	ctcatacaac	agggacaatg
94741	agcactctta	tttcaaaggc	tcattgtgag	gaacaaaatg	aaataatgtc	ggctgggtgc
94801	agtggctcat	gcctgtaatc	ccagcactct	gggaagctga	gatgattgga	tcattttagag
94861	tcaggagttc	gagatcagcc	tgcccaatat	ggtgaaagcc	cgtctctact	aaaaatacaa
94921	aaagaggcca	ggcgcaggca	gaccacgagg	tcaggagtgt	gaggccagca	tggccaacat
94981	ggtgaaaccc	tgtctttact	aaagatacaa	aaaattagct	gggtgtgatg	gtgcacacct
95041	gtaatcccag	ctacttggga	ggctggggca	ggagaatcgc	ttgaacccag	gaggcgaggt
95101	tgcagtgcg	tgagattttg	ccattgcact	ccagcctggg	cgacaggcg	aggctccatt
95161	tcaaaaaaaaa	aaaaaRaaaa	aaggctatgg	ctcaYgcctg	taatcccagc	actttgagag
95221	gccaaggtgg	gtggatcatc	tgaaggtcag	aagttcaaga	ccagcctggc	caacatgggtg
95281	aaaccccatc	tctacaaaaa	atacaaaaaa	ttagccgggc	ttgctgggtg	gcgctgttaa
95341	tcccagctgc	ttgggacact	gaggtgggaa	gatcgcttga	actcaggagg	tggaggttgt
95401	agtgcagcaa	gattgcacca	ttgtcccagc	ctgggcaaca	agagcaaaat	ttcatctcaa
95461	aaaacaaaaa	aaattagctg	ggaatggtag	caggcacatg	tagtctcagc	cacttgggag
95521	gctgacgtag	gagaatcact	taagcccaga	aggcagaggt	tgtgggtgagc	tgagatctcc
95581	ccactgcact	ccagcctggg	cgacagagtg	agagtctatc	tcagaagaaa	aaaaaaaaaa
95641	catgtctgtg	ccgtgtctgc	cacttggaa	aaatggcagc	tcactagtcc	ccgagggtaa
95701	gggacctcac	tttgccttgc	ctgtttacta	ctctgtaaaa	ttgaaataac	gcagctgggt
95761	gcagtggctc	acgcctgtaa	tctcagcact	ttgggagtg	gaggctgggt	gatcacctga
95821	ggtcaggagt	ttgagaacag	cctggccaac	atggcgaaac	tccattgcta	ctaaacatac
95881	aaaaattacc	tggatgtggc	ggcatgtgcc	tgtagtccca	gctacttggg	aggctgaggg
95941	aggagaattg	cttgaacctg	ggaggcggag	gttgcagtga	gccgagatgg	tgccactgta
96001	ctccagcctg	agcaacagag	tgagactccg	tctcaaaaaa	aaaaaaaaaa	aaaaaagaaa
96061	taacactcta	tgtaagggtt	actgaaagga	ttaaatgagg	ccatataaac	atataggcac
96121	agtgtcttgt	ggaggggggt	ctgggtaaag	cactggctcc	ttcctccctt	ctcagctaca
96181	gggcagggcc	ttgtctgccc	tgggtgtaggt	catggaccgt	cttgggtccc	ccttccattg
96241	tcaggctgta	ctgggttctt	cccacaggg	ctggcctagg	agcaggatgt	gttgtgagt
96301	tgggtgtgact	cttagtctct	gctttctaag	gacttgaaaa	atagaggaa	ccaggtgggc
96361	atacttggac	agtgttgccc	tgttgaccac	ttctcgctga	acttgggaag	gagggaatgt
96421	tttctggatg	ggaagaatcc	agtcaggcct	cctggagggg	gtggtaagaa	tctcaagcct
96481	tgaaggatgg	gaaagagcac	aataacaaca	gtgacaaagg	atagatcaca	aagcatcatg
96541	gactcacctg	actcatgtc	cgtgccaggc	tgaggaaagt	ttttctgctg	gcagtgcag
96601	aaggaagggc	attctccagg	tagaaggagc	tgtttgtgca	gaagcatgaa	ggccgctaaa
96661	agggggctgc	tgatagcaat	gaggaggctg	gggctggagg	cagggaagtct	gaggcgccct
96721	agtaaattgt	ccagacctga	gatggggcca	tgactacctg	gagcaggaca	gagaggagaa
96781	ttccagagag	ggcccaattc	cagagatgtc	agaagtagac	tYgacatcca	ctgtattagg
96841	acttggccag	gccgagggag	agagagtagt	ctgaggcaag	cctagtattg	ggtcttggac
96901	actgggacca	aggggtgatgt	tccagataaa	atgagcaacc	caggaggagg	cgagggtttg
96961	caggaggagg	tgatgaggtc	ctttctgcac	atcttgagtc	cgctgtccct	gtgggagggg
97021	cctctggggc	agtcgcatgg	gcaggcaggg			

chrom 4 genomic sequence (SEQ ID NO: 2)

>4:36870401-36969350

1	tacaaaagca	tttccataaa	tatatacaat	cctgtcttgt	caatcaaaaa	aaaaaatgaa
61	aaaataaatc	caactaaaag	taagatatta	acataatgtg	ctaaattaa	caaagctgaa
121	ttatctatct	actcatggta	atgttcattt	gttaaaataga	ctaatctgtg	
181	ataattcctc	ataacataag	aagttgatgc	Mtttacaaaa	acttatctgg	gatagtttat

241	tgacttgact	gggtcttaca	aatatagatt	aggtatatgg	atagaaaagg	atttgactgt
301	aaaatttgca	aaaacagtgt	ctcaaagcat	agtgatactt	taacatgatg	gaatttttgg
361	ctatgggata	gttataactc	acttctggca	atgtgagctg	actttatttg	tgttaaaatg
421	tatcaatcat	cttataaata	aataacagag	ctattaataa	caattttgaa	catttttccc
481	tacttgacct	tttaaaaaatc	agctccataa	gagacaaatc	aattaatatg	tgtttccact
541	attttttttc	ttacattaca	ctttgctggt	gactttaata	aggcagagaa	attggtgagg
601	gatggcactg	gtattgggga	aggcgtcaca	gtgaagactg	tgccatttct	ttggcaggca
661	gattacatac	ctcacaaaac	tccctgggag	atcagtattt	taagccctgt	tttttaggct
721	atctcatcat	ggccaagggg	catagataac	taaacttagg	gtcaccatcc	agtctggcag
781	tagagttaca	gaaagggcct	ccaaggagtg	aaccattcag	agttgttctg	gagattacac
841	tgtgaaagaa	gtgtctctgg	ttcatagtgt	atgagtcctt	tctatgctaa	ctctacaaat
901	tcaggagggg	aacaaaaaaa	cagaacactt	ccttttcctg	ttaaagctca	tagtctttca
961	gattcccaac	ccatagttag	acttggaag	cattgttcac	tgccctgggga	gaggggtacta
1021	tcaatggtac	tactatcaat	ggtagggagg	ctggaatcca	gaaccctaact	tcagatccag
1081	cagatgaatc	gttgaacaaa	ttcagaaaag	tgaatcttga	acctcaataa	gtgaggggtcc
1141	cacctaacac	atcaagctta	agccacaaga	aatattccag	ccccaccatc	ttgatagacc
1201	tgagcctacc	tacctgtaaa	agccaggtct	gggttcacag	tgcttatttt	cctaccattg
1261	taccttgctt	gtatcacatt	gtgttttaaa	tttatgctga	ctagtccatc	tcccctacca
1321	gacaatacat	aaagatgggg	attgagtcct	ctcaccattt	ggtagagaag	gtaccatttc
1381	ctaccattgt	accttgcttg	tatcacattg	tgtttttaaa	ttatgctgac	tagtccattc
1441	cccctaccag	acagttacata	aagatgggga	ttgagtcctt	tcaccatttg	tattctacc
1501	atttggtaga	atgtctggta	aatgtagcac	tcacatgag	tttggtgaat	gaacaaataa
1561	acattgatag	ctttggccaa	aataaaactgc	aaatttttaa	ggatatgaac	ttcaaatat
1621	aagatcctac	actacagttt	tgaggatgt	agcaggcaaa	agtgactgat	aaatatgttc
1681	agttattatt	acactacacc	cagaccatga	gaaagcagtc	ttagctatat	tcattgtacac
1741	atagaatagt	ataggataat	aatgacctgt	acctataatt	atacctttat	gcctatacct
1801	aaatctctac	ctatttgttt	ctgtgtcatc	tgaatgaata	cacagctaga	tggtatagata
1861	gatgatagat	agatagatag	atagatagat	agatagatag	atagatagag	gcatgaggac
1921	caatttgaaa	atataagcac	tttatgagcc	acatgttttt	acattgtaat	catttggttt
1981	ctacaaaaaa	ttgaccccgag	agagggtaag	caacaacaca	tatcatactc	tttgttccag
2041	caattagagt	gctgtggttg	ttaagatatc	aattataaaa	tattttatta	aaacacaaac
2101	acttttttct	ttttgagcaa	gatgattgat	tattcacaaa	tagtttcaaa	ccacttgtca
2161	ataaaaaatga	gcacgttttc	caattcatca	aaattatagg	acttcattac	actaaaaaca
2221	ataaatcaca	catctaacag	caaaccgacc	cctaataacc	tttggtacct	actttctaac
2281	atttcccaga	gtgtgacatt	gaactctcat	gaggtaaagt	gctctatttt	ctagctgtct
2341	gtggcaaaat	ataacagagt	aacatttaatt	tctaaatgtg	atttttgaga	tgaaagggtga
2401	ttggtaaatt	ttgcaattat	cttttgaaa	gtaagcaaaa	tatactttgc	tctttagaaa
2461	agattccatt	tctatatcta	cattataaac	catatatcaa	gtgattttac	gagaaaaaaa
2521	gtaagatcaa	gattacatat	taaacaaga	gaaatggggg	tgaagtctta	gacttggaac
2581	ctggaaagat	gatgaaacag	gagagagctc	ttttttggag	attctcctcc	ctttcaatgt
2641	ccatccccca	gcttctagag	ctctccccca	ctctgggagc	attcatggtg	ctctccagac
2701	cccagtccaa	tctcatgcat	caatgtcaag	taactccacc	tgccaatctt	attataaaatc
2761	tgatttcaga	ttgaaagata	agacaactgt	tccagatgac	tacttctctt	tttttccctt
2821	cggcggtggc	gggggggctg	ccatatttaa	ctgggaagtct	ctttaataaa	agaagttatc
2881	tcatagtttt	ttttttgaaa	tttttatttt	tttaatttat	tttatttttt	attattatta
2941	tacttaagtt	ttagggtaca	tgcgcacaa	gtgcagggtta	gttacatatg	tatacatatg
3001	ccatgctggt	gtgctgcacc	catttagcat	taggtatata	tcctaattgct	gtccctcccc
3061	ctccccccca	ccccacaaca	gtccccagag	tgtgatgttc	cccttcctgt	gtccatgtgt
3121	tctcattggt	caattcccat	ctatgagtga	gaacatgtgg	tgtttggttt	tttgtccttg
3181	agatagttta	ctgagaatga	tgctttccaa	tttcatccat	gtccctacaa	aggatatgaa
3241	ctcatcattt	ttatggctgc	atagtattcc	atgggtgata	tgtgccacat	tttcttaatc
3301	cagtctatca	ttgttgagca	tttggttggt	ttccaagtct	ttgctattgt	gaatagtgcc
3361	gcaataaaca	tacgtgtgca	tgtgtcttta	tagcaacatg	atttatagtc	ctttgggtat
3421	ataccagta	atgggatggc	tggtgcaaat	ggtatttcta	gttctagatc	cctgaggaat
3481	gtccacactg	acttccacaa	gggttggaact	agtttacagt	cccaccaaca	gtgtaaaagt
3541	gttctctatt	ctccacatcc	tctccagcac	ctgttggttc	ctgacttttt	aatgattgcc
3601	attctaagtg	gtgtgagatg	gtatctcatt	gtggttttga	tttgcatttc	tctgatggcc
3661	agtgatgatg	agcatttttt	catgtatctt	ttggctgcat	aaatgtcttc	ttttgagaag
3721	tgtctgttca	tatcctttgc	ccactttttg	atggggttgt	ttgtttgttt	cttgtaaaat
3781	tgtttgagtt	cattgttagat	tctggctatt	agccctttgt	cagatgagta	gttgcgaaaa
3841	ttttctccca	ttttgtagg	tgctgtttca	ctctgagggt	agtttctttt	gctatgcaga
3901	agctctttag	tttaattaca	tcccatttgt	caattttggc	ttttgttgcc	attgcttttg
3961	gtgtttttag	catgaagtcc	ttgcccatgc	ctatgtcctg	catggatttg	cctaggtttt
4021	cttctagggt	ttttatgggt	ttacgtttta	gtcttttaac	catcttgaat	taatttttgg

```

4081   ataaggtgta aggaagggat ccagtttcag ctttctacat atggctagcc agttttccca
4141   gcaccattta ttaaataggg aatcctttcc ccattgcttg tttttctcag gtttgtcaaa
4201   gatcagatag ttgtagatat gcggcggtat ttctgagggc tctgttctgt tccattgatc
4261   tatactctctg ttttgggtacc agtaccatgc tgttttggtt actgtagcct tgtagtatag
4321   tttgaagtca ggtagcatga tgccttcagc tttgttcttt ttgttaggat tgacttgggtg
4381   atgtgggctc ttttttggtt ccataatgaac tttaaaagtag tttttttcca attctgtgaa
4441   gaaagtcatt ggtagcttga tggggatggc attgaatcta taaattacct tgggcagtat
4501   ggtcattttc acaatattga ttcttcttac ccattgagcat ggaatgttct tccatttggtt
4561   tgtatcccct tttatttcat tgagcagtggt tttgtagtct tccttgaaga ggtccttcac
4621   gtcccttgta agctggattc ctaagtattc tattctcttt gaagcaattg tgaatgggag
4681   ttcactcatg atttggtctt ctgtttgtct gttattgggtg tataagaatg atctgtattt
4741   ttgtacactg attttgtatc ctgagacttt gctgaagtgt cttatcagct taatgagatt
4801   ttgggctgag acaatggggt tttctagata tacaatcatg ccattctgcaa acagggacaa
4861   tttgacttcc tcttttctca attgaatacc ctttatttcc ttctcctgcc taattgccct
4921   ggccagaact tccaacacta tgttgaacat accagaatct ctgggacaca ttcaaagcag
4981   tgtgtagagg gaaatttata gcactaaatg cccacaagag aaagcaggaa atatccaaaa
5041   ttgacaccct aacatcacia ttaaaagaac tagaaaagca agagcaaaaca cattcaaaag
5101   ctagcagaag gcaagaaata actaaaatca gagcagaact gaaggaaata gagacaaaaa
5161   aacccttcaa aaaattaatg aatccaggat ctgggtttttt gaaaggatca acaaaattga
5221   tagactgcta gcaagactaa taaagaaaaa aagagagaag aatcaaatag acgcaataag
5281   aaatgataaa ggggatatca ccacaaatcc cacagaaata caaactacca tcagagaata
5341   ctacaacaa ctctacgcta ataaactaga aaatctagaa gaaatggata aattcctcaa
5401   cacatacact ctcccaagac taaaccagga aaacgttgaa tctctgaata gaccaataac
5461   aggatctgaa attgtggcaa taatcaacag cttaccaacc aaaaagagtc caggaccaga
5521   tggattcaca gccgaattct accagaggta taaggaggag ctggtaccat tccttctgaa
5581   actattccaa tcaatagaaa aagagggaaat cctccctaac tcattttatg aggcctgatt
5641   catcctgata ccaaagcctg gcagagacac aacccaaaaa gagaatttta gaccaatatac
5701   cttgatgaac attgatgcaa aaatcctcaa taaaatactg gcaaaccaaa tccagcagca
5761   catcaaaaag cttatgcacc atgatcaagt gggcttcatc cctgggatgc aaggctgggtt
5821   caataatatg aaatcaataa atgtaatcca gcatataaac agaaccaaag acaaaaacca
5881   tatgattatc tcaatagatg cagaaaaggc ctttgacaaa attcaacaac cgttcatgct
5941   aagaactttc aataaattag gtgttgatgg gacgtatctc aaaataataa gagctatcta
6001   tgacaaaccc acagccaata tcatactgaa taggcaaaaa ctggaagcat tccctttgaa
6061   aactggcaca agacagggat gccctctctc accactccag atgactactt ctcaaattgtg
6121   ccctaaacat gctaacatgt aagtgaacca taaataatat catgtttctc attggcattc
6181   ttatattcaa taaaataggt cattatatta catatgacat ctaatttgta aaaaatatac
6241   ctttccccc tgggactcag tgtaactatt tgcattctga gtgtggaaaa acctgtgtcc
6301   catgaaaacg tctgtattct ttaagtggct ctactccaa tgtagataga tataggcata
6361   ggtgcagatc tagatgcagg tataggatata ggtgttatt actgtacatc atgttatata
6421   aggaattctc ttctccctac tctttttctg aaacaaagct ctctttttca caggctactc
6481   gcactcagct tctagttatc agaaaacagt tagtaattta ttttaaaagt aacttcagga
6541   agaaatggat gttgtgaatg aaggaattta aagcagcatg aagaataaaa tcagtcattt
6601   agtctgacag gcagtcgtgg gcagctattt ggggtataag gcaactatct tgctgtttaa
6661   accaaacccc atgccacatg cagcaccatt ggagtattat gtaagtgttc aaccagtgcc
6721   aaaattctat cattcttagc tgtctactca ctataggtag ggccgatccc caaattctaa
6781   aagattgaaa tggataggat cttgatacca tttttcttgt atgttttaaa atagcaagcc
6841   actcacttac aaatttgaa aggagattgc atttgccaaa catgtagtta taattgagat
6901   tttcattgtg tacctctttg cgaaaatgca gatggattag cattcatggt cactgagttc
6961   tgtttccatg gatatagttt cagtcatgcc aaccagggac agatttgtta ctttattact
7021   ttgtttcaag tattattcaa acatttaaaa aatatcccag cactatacaa cattatatgt
7081   taatgtagt gtaattttt ggtaatttt cgttgacttt aaatgatttt gaggttccaa
7141   tcacgatttc ttaaaagtat gcacctagat tgtaatactg aaagcctatt ggtttaacta
7201   ggcattgctag cttggcRtcc agggaaactcc tgaagaagcc agaattaagt gaactcgcag
7261   tttcattaat atccaaagaa gagtttagcct ggtaacaaga tgtttcatct actacaattc
7321   caaaaagctt tagaaaaggc aatggaggac tacagctagg agaaaattta ctggtcagat
7381   aaatacaaaa attatttcac tcttccctac tagactagta aatttttaaa aactatgttt
7441   ttatattctc gtatttctca caaccagtt tttaaaattt attagtacct aatacacaat
7501   tgtaaaattta atgtcacata tctgttggtg gtcttttaat tagaacctgg taccacaga
7561   gttttaaata ctgacagtga tgagcctgga acgaataccg aatatgcatg ttattttcag
7621   gaagtgtgtg aggcattcca agtggcataa aaaataatgt gtacatagct tgaatacgtg
7681   gaagtttatt atataggat aagtttaattc aattcaataa gtaatttaat tcaatgcact
7741   ttggatacat tggagggaaa aacagtcaaa aaattttact catggatctt atatcctagc
7801   aaggagagacc accaataaat aatagatatg ataaacaatt gtgtgggtgtg ttagaagggtt
7861   acaaatgttt tggaaacaaa gaaaaagtag agggMaaagg gtaatcttga atagtagtgtg

```

7921	tgtgggctgt	tgccaagttg	cagtttttaa	aatgatcaaa	gtctgcctca	ttgacaaggt
7981	gagattttaa	caaaggcttc	aggagttgag	agagttggcc	aagcaggtgc	tggggaaaga
8041	gctttccagg	cagaggcaac	acctagagca	aaagtccaaa	ggtgagagcg	tgctgatgt
8101	gttcaaggaa	gaccagggac	ataagtgtag	ctctccaggg	tagtgcaagc	agggaagagc
8161	tgtagaaaat	aaggtcagag	aggtaatgag	gaagcacact	cagtcataca	cagcctgaca
8221	catggtgatt	atatgaactt	tagcttctgc	tgagtacagg	gaagagccat	tgcaagaattc
8281	tgagcagaga	aacagcttac	tgagatgtaa	gtttcaaatg	gatgattcta	tttttgtata
8341	gagtacagac	tatagagaga	agaggataaa	tgaggtgac	caatttggag	aggggtgcag
8401	taatccaggg	acaacagatg	ggggtcaga	tcagagtaag	aggtgctaaa	aagtagttaa
8461	attctggata	ttgtttgaag	gtagagaaaa	aagaacttct	tgatggattt	aataaggaaa
8521	acatgagaaa	ggggggaata	gaggatgact	ttggcctgtg	caactcaaaa	gatggaattt
8581	cctcaactga	aatgaggaaa	actgatgga	gaaggttcca	ggaagaagat	caggcggttca
8641	gtctaggata	cggtcaagca	gaattgttag	acagttgtgt	atatgaatct	gaagttcaag
8701	tacaaaggtc	tggagtgagg	gtttaaattt	gggagtcacc	agcagataga	tggtactcaa
8761	agccactggc	ctggatgatt	ccaacaagag	agttcatgtg	tggaccaagg	actgagcctt
8821	ggcaatacaa	ctttaagtca	ggacgaagag	ggagattaac	cggaaaggag	cgccagtgga
8881	gttagaagga	aagcaaaagga	tggtgtgggt	ttttttcaaa	agctaagaga	agaaagtgcga
8941	tcaaagggga	aggagaaatc	aatttgggtca	aaacctataa	taggtcaagt	gagatgaggg
9001	ggaaaaaaa	tggccgtttg	gtttagtgc	acagaaactag	ttcattggta	aatttaaata
9061	ttcaaccttg	catcagtatc	gccactttct	gaacactgcc	attttgattt	atgcattgaa
9121	ttcccccaa	caaagtttta	tgcttctggc	tttcttttat	ttttccttat	aactgtcctt
9181	ttctttgggtg	tttatgggtg	aatattgggt	ttctcttaaa	tcttttaact	ccttatattt
9241	tatgtgcttt	ctcggcacat	gcttttgaat	tacattttta	aattttaatt	tcatctcaca
9301	tttataatat	agagcaatgt	ttgaattcag	ctattaactt	ttatccatct	cttagtctat
9361	attgattttt	tggttgtctt	tcacaaatcac	attcatgcct	tatttttagct	gccacattgt
9421	ataaagttaa	ctagcattgt	ttaaactttca	catctagtaa	gtagcatagt	aactgtattt
9481	taactgtaat	tgacaataaa	catctgtttt	atctcccat	tttttgttaa	ttcaacaaca
9541	cagttgtatt	ctacctgggtg	cattcattga	atcttgctaa	aaaaaaatct	aatcacacc
9601	aatattttgt	ttctaactcc	tcctcttcaa	cagcctctag	attatatgca	ttttccatta
9661	cttttttttg	tattgcaaat	tactgtaatc	ttacttttgt	tacttttatt	tctggagatt
9721	cttttggttac	tgttttacc	ctgtgacttt	catctcgct	gacctttgt	atcttctatt
9781	tattcagatt	tgtgtcatct	tgctattgta	ggtacattta	tacatcctgg	aatggcattc
9841	cagtttactc	agtgttctca	aggcaataaa	tgctatgtcc	tgatgtgaaa	atataactat
9901	taatataaaa	tgagtgtgca	ttagagactg	caatcagtag	cagaatagag	gtaatatata
9961	ttgatagtta	ggaatattga	aaaggaagtt	aaaaggggaa	aattttattag	ctagctagta
10021	ataaaactct	agaaaggcca	atgcaaaagac	tttctatcta	tgcaagtcca	gagaactata
10081	catctaagtc	atagctctac	acaattgtca	aatccttccc	atttaaatag	taaatacatg
10141	cccaacatac	atacacttaa	acatcattaa	acaagaataa	gccaaggagc	atattaatag
10201	tcatcttcat	ctcttctctt	aatacttgct	ctgcttatat	ggacagacat	tttgagttct
10261	gacagaccat	ttatttagat	aggttcaatt	agcaagttat	gcaaggtagt	caaactgatt
10321	tgaaaaaaa	atgaatgtca	ttgaacactt	cactgcacag	gatgtcttca	gttttttatt
10381	tccaataaaa	tagtaagcta	tgtagtagac	tcttatggta	ataatttctg	tattctaaaa
10441	tgtcaacaaa	agaaaaatat	ccctcagaat	ttacttatgc	ttcagagatc	tttctaagga
10501	tttaaaacaa	tttaggtagc	cactctttta	aaaagaaaac	aaaacaacaa	tatgaaactg
10561	tacatatgta	attctctgtg	actcaaaact	caaaagtgtt	ctttccacgg	tagaaactca
10621	gaaggtagag	agaaaaacca	gctattttat	gtgtatgttg	ttttcagtta	ttatcacatt
10681	cttaatgaat	tttctggagg	tctacaatga	ccctagatcc	catgaaatga	taaattaaat
10741	gtgaaagtgt	gtcaagtact	tgtataagat	actccaatgt	cttcagaatt	tacaaaattt
10801	tctctgaagt	tatcaaggca	gaaaagtgca	gcatcattg	agtcacgaga	tataaattct
10861	cttcatagct	ggttcaacta	gctgattgat	ccaatgagtg	ttgctctgcc	tccccattat
10921	ccattcaccc	ttatcctgag	agtcacttcc	tgatttccct	atcgggaaag	atgcctcaac
10981	tctgaatttt	ataattgtta	tgctttatgt	tgcatagtgga	tgaatatata	aacctactta
11041	aaagaagtgg	gaattttatt	ttttatctag	agtctttgtt	caggcaatcc	ttcaccagga
11101	gttccaaaaa	tgggaaaaaa	gacctagctt	ctatatcaat	tttttagttg	cttttatctc
11161	tgtgttatct	ttacttgatt	tcagctagtg	aacctaaaaa	cccttaacca	aacatacaaa
11221	tcaagatcaa	aaaacaaaaa	tgtagagaca	taacatgaat	ggcagctgaa	ttagtttctc
11281	gtggctgcca	taacaaatta	ctacaatctg	ggaggcttaa	aagaacagaa	atttatctct
11341	tcacaattct	taaatctaca	aatctaaaaa	gaaggtgtca	gcagggtgtg	gcttcccttg
11401	aaggctctag	agaaggatct	tttcttgctt	tttccagctt	ctctttttcc	tataaagatg
11461	ccagtcattg	gattgaaggt	ccactctagt	ccagtaaaat	ctcatctcaa	atcttacctt
11521	attacatctg	caaagacact	acttcaaaa	aaggtcacat	tctgaggttc	tggatgggca
11581	tgaatttttg	gggaacatta	ttcaatccac	tatagagata	aataacaata	aacaccaatt
11641	gcattgtgtg	actctcagga	cctcattaga	tctctttgga	tgtagttttt	cttttctctg
11701	aaaatatttg	aaatgggcca	aacaagggaa	tctttaaggt	ttattctatc	cttgccttca

11761	atgaatccta	ggatcaaagg	acaatgtcat	ccaagtgttc	aggctatgag	cacaaaaaat
11821	tgggcttttc	tagtcataca	tgtacttata	tacaagggtc	acatattatg	agtttggtct
11881	tgtaggacat	tttttcattc	tggagacatg	aaagatttct	tatcaagtct	agctgaattt
11941	gggcaatgtg	tgactcctga	ccatttttca	gtgatgttct	tggccatgat	attcagaaga
12001	ctgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	gtatcccaag	tagctggaac
12061	tacaggcgcc	cgccaccacg	cctggctatt	tttttgtgct	gttagtagag	atggggcttc
12121	accatgttag	ccaggatggt	ctcaatctcc	tgacctcgtg	atccaccgcg	ctggggctcc
12181	caaagtgtctg	ggattacagg	tgtgagccac	catgcctggc	ctatatatttc	ttttccaaag
12241	attttttaag	accatttgct	cttgtcttct	ccctgttaag	gattatagct	ttttaaatag
12301	cataatgata	acttagcaac	taaagaattc	cacacagatt	tataagatga	atctgtttac
12361	cagatatatt	aatctaactt	gaaaacattt	tttagtactt	tatttgcagg	atgcttttta
12421	gcataaatgt	tgttggaatg	aatttggcca	tttcaacagt	ctttatttta	atttgtttta
12481	taactaaaat	aatgttaatg	agacggtggt	cttttccttc	aaatagaatg	ataggcaagt
12541	gcaagaatgt	tagagatcaa	cttgggtggt	gatgtttctc	ttttgagaac	atggacatgg
12601	gttgtgagtc	tgttctcatgt	aatcccaggc	actagggagg	ctgaggcagg	agaatcgctt
12661	gaaccagaaa	ggtggagggt	gcagtgaagt	gcactccagc	ctgggggaca	gagcaagact
12721	cagtctcaaa	aaaaaaaaaa	gaagaagaag	aaaaaagaaa	agaaatgtaa	cattgtaaaa
12781	tgatagtccc	tatggaaggc	agtatgacag	ctcctcaaaa	aagtaaacad	agaattaccg
12841	tattgatcca	gtaattccat	tactgggtgt	ataccaaaaa	gaactgaaag	caaagaatca
12901	agcatatatt	tgtactccta	tgtttatagc	agcattattc	acaacagcca	aaaagtggaa
12961	gcaacctaaq	tgtccactga	aaggtcaata	aataaacaaa	tgtatgatat	acatacaatg
13021	gaactttaat	cagcctctga	aaaggacatt	ctgatacatg	gtacaacata	cataaacctc
13081	gaagacatta	tgttaagtga	aataagtcaa	tcacaaattg	acaaatattt	attattccag
13141	ctctacgagg	tctctagaat	aatcaaat	atagaaacag	aaattagaat	gatgggtgta
13201	atgggctaga	gaaatgggaa	aatggggagt	tagtatttaa	tgggtacaga	gtttcactat
13261	gggaagaaga	aaaatcctgg	agatgtagtg	tgggtgatgg	tgcacaaYga	tgtgaacata
13321	cttaatgaca	ctgaactgta	cacctaaaaa	tgggtaaaaa	tctaaatatt	aattatgaag
13381	tgagcattaa	ctaacatagg	ccatattcta	gagtattaaa	acttaactgg	aacaggaggc
13441	caggcacagt	ggctcacacc	tgtaatccca	gcactttggg	atgctgaggc	ggatggatca
13501	cctgaggcca	ggagtctcag	accagcctgg	ccaacatggt	gaaactctat	ctctactaaa
13561	aatacaaaaa	aatctagctg	ggcatggtgg	cacatgtcta	taatcccacc	tctctaggag
13621	gctgagtcag	gagaatcact	tgaaccaggg	aggcagaggt	tacagtgagc	cgagattgtg
13681	ccactgcact	ccagcctggg	caacaagagc	gaaactccgt	ctcaaagaaa	aaaaaaaaaa
13741	aacttaatta	gaacatactt	aatgacactg	aactgtacac	ttaaaaaatg	ttaaaattct
13801	aaatattaag	tatgaaatga	gcattaacta	agacaggcca	tattctagag	gattaaaaaa
13861	gtcacaaaaa	attttaaatg	gtataattca	tatgaattat	tttctctgac	tataataaga
13921	ttaaatagg	attctagaac	aatatctaga	aaatccctca	aatattttga	aactaaatac
13981	acaagtctta	gtattttatg	catcaaagaa	gaaatcataa	atgaaaataa	gaaagcaaata
14041	tcatctggat	gaaaattaaa	acatctcaaa	atctttgaga	agtagctagg	gtagtattta
14101	agtggaat	catagcccta	agtgtcaata	ttataaaaaga	agataaattg	caagccaatg
14161	gcctcaggtt	ctacttaaga	aaaagagaaa	caaatataaaa	ctgaaataaa	gtaagggaat
14221	attaaacatt	ggggtagaaa	tcaataaaat	ataaagcagg	aaaacaatga	agaaaaaaMt
14281	gaaatcaaaa	gctggttctc	tgagacattc	aatgaactgt	taaacctcta	gctagactta
14341	ttaggaaaaa	aatagaaaag	acacaaaatt	ctaatagcaa	gaatgagaga	gggtgtacaa
14401	ttagatagc	tctattgaca	tttaaaaaaa	taagagaatt	ctatttgcaa	attatgtaaa
14461	atatgtgaca	atcagaaga	aatatataaa	tttcttaaaa	gacacaaaca	actaatgaag
14521	ttcacttaat	aagaaataat	aacctgaata	acgagatatc	tattagggaa	attgaatttt
14581	aaacctcccc	ttaaagaaaa	ttcaaggccc	acatgtcttc	actgaggaaat	ttgactaagc
14641	athtagtgga	gaaataatac	taattctggt	caaatacttc	taagaaaatg	aagagaaaat
14701	tcatatcaac	tcattttaca	aggccagctt	atccctgcta	ccaagtgaag	acatttgaag
14761	aaaattacag	accaatattc	ctcaacgtgg	aggaaaactt	tcttgtttta	ttagtctga
14821	gagtttattt	gaagtaagtc	acaacaaaaa	ataataaata	cttgtgttgt	cttgggttta
14881	ctttcagcct	gaagtgtctc	aagcaagaaa	aacctgaga	tagaataaag	ctggatacat
14941	acggacatag	aaactggtat	tctcctttac	atctttccct	ttctttcaat	atctttttct
15001	ctctttattt	taactattat	tttaaggttc	gggatacatg	tgaaggatgt	gcaggtttgt
15061	tacataggta	agcttgtgtc	atggggattc	attgtacaga	ttacttcacc	accaggatg
15121	taggcctagt	acccattagt	tatttttctc	gatcctctcc	ctcctccac	cttccaacct
15181	ccaataggac	ccagtgtgtg	ttgctccctc	ctatgtgtcc	atatgttctc	atcgcttagc
15241	tcccagttat	aagtaagaac	atgtggcatt	ttgttttctg	ttcctgcgtt	agtttggtta
15301	agataatgtc	tttcagctcc	atccatgtcg	ctgcaagac	atgatctgt	ttttctttat
15361	gactgtatag	tattccatgg	tatatatgta	ccacatcttc	tttatccagt	ccatcatgta
15421	tgggcattta	ggttgattcc	ctgtctttac	tactttgaat	agtgtctgca	tgaacataga
15481	cgtgcatgtg	tctggaagtt	gaatcccaca	atatagaaaa	aggtaacaaa	aataaaagt
15541	gggtctatcc	caggaatgca	agtttagttg	aagattcaag	aatcaattaa	tgtaatctat

15601	tataataata	gactgaagaa	aaaaacataa	tatctttacca	atacacgcag	aaaagtgggt
15661	gacatatcca	acatccatcc	ttgatttaag	ttaaaactga	aaactctcag	caagcttaaa
15721	atagacggga	acttccctaa	cctaacaaaa	ggcatgtagg	taaaacctac	agctggcaca
15781	cttagtagtg	aaaaattaaa	tgctttcccc	taagttcagg	aaccagaaaa	ggataccac
15841	tctcatgatt	tctattcagc	gtcgtattag	atgttctagc	caatgcacga	agctgagtac
15901	aataaataaa	agacttccat	attagaagaa	aaactatgtt	tattcacaaa	taacatgagc
15961	atctgtgtcg	aaaagtcaac	ggaataataa	aagttactag	aactaataag	agtttttcag
16021	agtctaaaga	tttgaaacta	gtatattaaa	attaattgta	tttgtatata	catgcataaa
16081	tctcagaaat	tgaaacttaa	aaaacaagta	tttaaaactag	cagcatagaa	tgtaaaatag
16141	ttagggataa	atctcatgaa	aatgtgtaag	atctatacag	tgaaaattgc	aatacattta
16201	ttttttaatt	attttatttt	attattatta	tacttttaag	ttttagggtt	catgtgcaca
16261	atgtgcagg	ttgttacata	tgtatacatg	tgccatgttg	gtgtgctgca	cccattaact
16321	cgctcatttag	cattaggtat	atctcctatg	gctatccctc	ccccctcccc	cacccccaca
16381	cagtcocccg	agtgtaatgg	tccccttccg	gtgtcccaag	gttctcattg	ttcaattccc
16441	acctatgagt	gagaacatgc	agtgtttggt	ttttgtcctt	tgcaatagtt	tgctgagaat
16501	gatggcttcc	agtttcatcc	atgtcccctac	aaaggacaag	aattcatcat	tttttatggc
16561	tgcatagtat	tcctatggtg	atatgtgcca	cattttctta	atccagtcta	tcattgttgg
16621	acatttgggt	tggttccaag	tctttgctgc	aatacattgt	taagagagat	taagtaagac
16681	ctaaatcgat	ggaagataac	atcagttcca	tgagttggaa	gaattgtaat	atgtcaattc
16741	tgtgtacatt	aaccacaaat	tcaatgcaaa	ccaggcaaaa	tccaagtagg	cagattggta
16801	gatatacaaaa	acctgaatct	aaaattgggtg	cagaaatcca	aatgtcctaa	tgtagcctaa
16861	acaactttga	aaaagaaaaa	gaaattggaa	agacctctag	taagaacaac	atgaacaaat
16921	gtcaaagtga	aagaaactag	aaacaaaaga	ctacatgtta	taagatttca	tttatttaac
16981	agactttaaa	aaaacagaaa	ctatagtcat	ataaagcagg	tcagtRgttg	cacggggtct
17041	aggcatttgg	aggagagaat	taattgcaaa	gggacacaaa	tgtatgttta	aaattttaca
17101	tacagtaatg	tgatgggtgg	ttacaccact	gtctatatatt	gtcaaagctc	atcaaattat
17161	atattttaa	taatgaattt	tattatacag	tattatacta	gagctaacca	aaataaaagt
17221	gaaaatctct	gtgggggaaa	ctatagagaa	acataactc	ttatatgttg	gtggtaagat
17281	ttcaggactg	ttcaaccact	atggagagaa	ggtagtaaat	atctagtaaa	atcacacatg
17341	catatgttct	ctgactcaga	gtttctcctt	ttaggaacct	attccaatga	tacattgaca
17401	aaaatatgaa	atgggtgtaga	cacacagcta	ttcattgcca	acagtccttt	atagtagtct
17461	tataccataa	attaaagaat	aatgaataaa	caaagaataa	tacagccata	acaaaaagac
17521	acaacaaaaa	gtagtccatt	cagggtgggtc	atactggcca	aagatggaac	aatctgagca
17581	tcgattaaaa	taacggctgc	agtgaattga	aacctatcag	ttttgttaaa	ctccatgagg
17641	tcaaaatgat	acttgaaatg	tagctaactg	gtcatctcag	agagatgcta	tgagagcaac
17701	tcattattct	gaaaactgat	aaataaaagg	aaaagtacag	catatttccc	cccagttctg
17761	aatatactat	aacttgagg	aattgaatag	ttgtggaaaa	ttatttttaga	aaaattacac
17821	ctaagaaata	cagaaaagta	tgatggaatt	ataatatgac	tatttttaaac	ccctcctgaa
17881	aaaataaatt	tattcaattta	tcattgtggt	tgctaaaata	attaggtgga	gccaggcaca
17941	gtggcagggtg	ccagtcgtcc	cagttactct	ggagactgag	gcaggaggat	cacttgagtc
18001	caggagttca	aatccaacct	ggacaatata	gtaagacccc	catctctaaa	aagtaaaaaat
18061	aaaattatta	tgggaaaattc	tgatgcagat	tataacaatg	gatagaccat	agtgaataca
18121	cctaaatcca	ctgatcaacc	tgagcattat	taaaataagc	aatctgacat	tgtgggtctc
18181	ttgtttataat	gcagtaggag	tatttagcat	cacctacaaa	atctccttgc	caaattgtcaa
18241	acctaaatct	cctcaagcca	ataggtctta	Yaaactttca	caagataata	agagaaacag
18301	taaattatgt	taaacaaaaa	catagagaag	tggtcagaaa	actgagaatg	ggagaaattc
18361	tatagaaaaa	tggtcaattt	ttacacaaac	aaaaatgccg	agagagagaa	agagatgacc
18421	atagggttaaa	aagctgaaga	aaactttaat	ttcaagttag	ggtagtttgg	agccgaccat
18481	tgctcttgct	gagaaaaatc	aaaacatttg	gatcactctc	tctctctttt	tttttttttt
18541	tttttaattgt	taaaacatca	gagaactgct	aaagtaaccc	agactggaag	aaacaatggg
18601	agagaagtca	gaaaacatgt	ttcagccatt	aattccctgg	ggaaatttaa	tggttctgcg
18661	tatagagaaa	gagccttaga	atctaggtta	cgcccaggca	aagagacact	gaggagataa
18721	gagaaaccag	cagagcattt	aacagttttg	tcaggctaga	gagacaaaaa	tgagacataa
18781	aaggagacag	atccaggcag	gcagtgggtg	aaggattaaa	cctgaaacat	ggcaagtttt
18841	tctttcaaga	catttgggtg	attctaaata	tgtacatggt	aggaaatgag	aagttaggca
18901	gaaaatctct	gaaaagcaga	gtaggttttc	caacaatctc	agtaattgtg	gaattacagt
18961	ttaaaaccta	ctagatggag	gggttcccaa	caaacaccac	tgaacctcag	ttggaaacct
19021	gggaggcctt	tatactacaa	tcaaggtgga	ctgagaggtg	aaaaaaaaac	acccaaactg
19081	aaaccagta	ttgactcatc	tcacactttg	aattaatgtg	atcaggccca	cattgtcagc
19141	caagcagaga	aaagactaaa	ttctttcttg	aagaaaaata	catcatctgg	agtcataaac
19201	attactagc	tcaccaaag	aaaagagagt	gtaaaataaa	gtcaagaaaa	ataacatgat
19261	agaaatacac	caacaagtga	tcccaacatt	ggagttatta	gtaaattact	ttaaaatgac
19321	cttgaatgat	atctttaaga	aaatatacat	aaattttacca	gagcttttgg	aatctatttt
19381	ttaaaaaaca	aaggtaaatt	ctaaaactga	aaaataataa	taactaacat	taggaactca

19441	atacgtggctc	tgagcagatc	agacacagaa	gatgaaattc	tgaaactaga	agaaagagca
19501	atagaaaaca	ggtgacatga	ttttttttgt	tcttgttggt	gttttgatag	agacagggtt
19561	tctgcatggt	gccaggctg	gtcttgaact	tatggactca	agcaatcctc	ccactttggc
19621	ctcccaaagt	gctgggatta	caggcgtgag	ccaccacacc	togtccagaa	aacagggtgac
19681	gtgaaagagg	aaacaactaa	cagaaaatag	agaaaaactat	aagatgggta	caggatgaag
19741	tgaaatggca	aaatataaat	tttcttagaa	tcacagaagc	agagaacaga	aagaagcgca
19801	aaagcaaaac	tcgaaaagat	attagccaac	tctgatacat	agaaggaaaa	ccatactgta
19861	tttaggcaca	tcatagtaaa	actgcttaaa	accaaggaca	aaaagggaaa	tattaaattt
19921	acttagagaa	agaaagaaat	gttaaccaca	aagcaataac	aataagactg	agctattact
19981	ctcgacagaa	attgtgaaat	ttaaagcaca	atgatacaac	tgctaaaaga	aatgatggc
20041	caccttagaa	ttctatcccc	agtgagaata	gcattacttt	ttttttttga	gacggagtat
20101	ctctttgtca	cccaggctag	agtgacagtg	cacgatctca	tctcactgca	acctcttct
20161	cctgggttca	agcaattctc	ctgcaacagc	ctcctgagta	gctgggatta	caggcatcca
20221	ctaccatgcc	tggtctaattt	ttgtattttt	cttttagcaga	gatgggggtt	ccccatgtac
20281	cttttttttt	tttttagtaga	gacagggttt	ccccatgttg	gccatgctgg	tctcaaacac
20341	ctgacctcag	gtgatccgcc	tgcttgggcc	tcccaaagt	tataggcgtg	ggcaatcgca
20401	cctggccgaa	aataacRttt	tttaaacgt	aaaataaagg	cacgacactt	tctgacaaat
20461	actgaggtaa	tttgtcgcca	gcagaccttt	actaaataaa	cggatttgtt	caaccagaag
20521	aaaatgatcc	catatggaaa	cacacatgtg	tggaaaagaa	tgaaaaaacac	agaatatgta
20581	aatacattag	tatatgtaaa	gtaatatgtg	cagtacaaat	agtggaata	atatattgtg
20641	aaacttaaaa	tatttataga	attaaaaatc	atgggtggcag	aaatggaaa	acggtaggag
20701	gtaaatgggg	ttaaaatgtg	caaagggttt	aacatttagag	aagtggaaaa	tagtaatttg
20761	cattaaacta	gggtaagtga	aggatgccag	ctataatgtc	taataaccac	taaaagacta
20821	ttaaaagaag	taattaagta	aatgtagcaa	acatatttga	ttacttcaaa	agaaatcaag
20881	acagagaaaa	agcaaaaacaa	agcaggtagt	ccaaatagaa	aataagacta	cagacataaa
20941	agcaagtata	ccaataatta	cattaaatgc	catgaactaa	atgattctat	ttaacaacag
21001	tacttgttag	gttgaaattga	aaacaaaaata	caactctatg	ttgctaagaa	gaaatatact
21061	ataaaaaacat	aaattcgaaa	aagttaaaag	taaaaagata	cagcaagaaa	acaggacttc
21121	actcctacaa	cttgcaagga	actgaatttt	gccaacaacc	atgtaaactt	caaaaagaac
21181	tccaagtatc	tgatgagaa	atagctctgg	ccacaccttg	atctcacctt	gcacagagga
21241	caccatgcct	ggactcctaa	cctacacaac	tgtgagctaa	tacatgagtg	tactttttag
21301	ctgctgtatt	tgtggcaatt	tgttacacag	caatagaaaa	ctaatacgta	acatgaatgt
21361	ccatccaaaa	attaattaat	tagttgtgat	atattcataa	tatgaagttc	tgacagcaa
21421	tgagaataaa	tgcatatttg	ctatatgcat	tacctggaa	gagtcctcagt	aacataatgc
21481	agaacaaaaat	aaattagaca	caaaagagta	cctgcatgaa	tctatgcatt	tctagaaagt
21541	gcgtttgttg	ggttgtagta	aaatctcata	ttttctttga	taaaactcta	tatctttcat
21601	aagatttctca	agcaaatcat	gaccccaaaa	aggcaagggt	cactgatgta	gattatgtta
21661	actatgtact	taaaaaaaaat	taaatcagca	taattgcctc	aattttttatt	gttgatgtag
21721	catgaggaca	gaatgcaatt	tgtattgttt	ggtatctttt	aagtttttct	tctgctaatt
21781	atcatcctta	aatcctataa	aaagcaacaa	agaaacttaa	tgctatggaa	ttacaacttg
21841	agYgcaagt	taaatctttt	gcagggtttta	tcatttttgt	cttcaagaaa	ggaatccaag
21901	gctaggcatt	gtggctcagg	cctgtaatcc	cagcactttg	agagggtcaag	gtgggaagaa
21961	cacttgagat	caggaattca	ggcccagccg	gggcaacata	gcaagactac	atctctaaat
22021	aaaacaaaca	attagcaggg	cagagtgatg	tgtacatgca	gtcccagcYt	ctaggaaggc
22081	tgaggcagga	tgatcactcg	gactcagaaa	ttcaagtctg	cagtcagcta	tgatcacacc
22141	actgYacccc	agcctgaaca	acagagtaag	atcctgtgtc	ttgaaaagaa	aagggaagc
22201	aagaaagaaa	tccaatgttc	tgcaattctt	tcttattccc	tcccagtcca	atcatgaatg
22261	tgctgtctgg	ggttggtcag	attcgaggac	accaactagg	ccttctcact	gatgggtaga
22321	accaatgcaa	gtggtaagag	cttgagcttt	tcctctcact	ccctttgtct	tctcttttcc
22381	catgtgacag	caaccaagga	ttgggtcatt	tcaggagaat	tatttctagc	agaggtcagt
22441	atcagacatc	gcgaacaaat	atattctcat	tatgtactat	ctttctctct	ggtatcctct
22501	ctctgatatt	gacagatgYc	accagttact	agctagctYa	aaaagggtatg	ttctatataa
22561	catgccctgg	gttactaaca	tatactgcac	caaaatttgt	gccttcaatc	cagaagttaa
22621	taaagtgtta	ccttggtttt	cttcattcat	tgtttcttgc	tcatgcattt	tcctccagag
22681	ccttgtcaga	gaacatgggg	atgggtcagg	acaggatttt	ggtatttctg	gatcccagg
22741	tcttaaaactg	cttctatact	gtaaagaaat	ttcaccctat	tttctgaaag	ctgtaggtaa
22801	tacccaacct	ttgcgtatta	agacatacag	tcttgactta	aaacataattt	atttagggat
22861	tgattttctca	ctcaaatagt	ttaggaacat	tctcacgaat	tcacttttgg	aatgagaata
22921	tcatatttat	gttacactct	ctctgaagaa	attgcatata	ttcacatagt	gctaacatga
22981	cacaggaact	atgtttgaat	gtgtttgtag	aaaatagctt	tgacatcaaa	acttgctctt
23041	taaagccagg	aaattcatat	tcatacttag	aacatcatag	atggatataa	ggacaggtgg
23101	atagaaaatg	ctaaaaagt	ctttacagag	agcagagagt	gtgtttgaga	aatgtgaatc
23161	atTTTTTTTaa	atgtaaaaac	caaataacca	ttgccttagt	ttcaacttta	gcctctgggc
23221	aattatcaag	cctttKcatt	ttcctgagtt	cctttKactt	tgcagggttt	cacattccat

23281	ttactgcccc	aaacacatca	gcgaactggt	taattcaatg	atataagttg	tatttttcta
23341	ttaaaataaa	cacctttgtc	attgatgact	aattcaggta	agaaaagcct	tttggtttgg
23401	gRtaaaactc	attaattacc	taatgacaat	gaaattcagt	cattagcagg	gaaaggaaat
23461	atttctaggt	gttagagaaa	ttcattttta	ataacttaKg	ctgttttttt	aggataacta
23521	cctttcctat	tctgtaatac	acacaaaaaa	tggatttttt	cctgttacac	accattcaac
23581	acaaaaccac	acacaaaagc	aatcactttc	tcaacaaacR	tgaatgctgg	gtgtcctttg
23641	attccattca	attctgactc	atctacctgg	agatagcagt	ttgagggtct	agtcaccacg
23701	gactgccgtt	tcagatgcca	aatgcaagta	ataggttatc	acctgtagtt	ctaaccactc
23761	agctgtgaat	cagagttttt	atgaccctct	cctcatgttt	gattaattttg	ctagagtggg
23821	tcacagaagt	cagagaaata	tttatgttta	cccgtttatt	ataaagaaca	Wtggaaaggt
23881	tacagataaa	agtccagatt	gaagagatgc	atagggtgag	gtaagggaga	aggataacag
23941	agcttcaaa	ccatctccag	gaacctccac	atagttagct	gtccagaagc	tccccaaacc
24001	aagatatttt	tagtttttag	gaaggcttca	ttacacaagc	gtgatggata	gaattattgg
24061	ccactggcga	tcaattcaac	ctttagctct	tctccccctc	ccagagggtt	agatgtgggg
24121	ctcaaagttc	caactStcta	atcatgccct	ggcttttctg	gtcaccagcc	ccaatcctga
24181	agctaccta	gggttcccaa	acctagtcat	ctcattagca	tgactaaagg	cactcttaga
24241	cactctgcca	gggccaggca	cagtggctca	catctgtaat	cccagcactt	tgggaggccg
24301	aggcggatgg	atcacttgag	gtcaggagtt	tgggactagc	cttgccagca	tgggtgaaacc
24361	ccatctctac	taaaactgca	aaaattagcc	aggcgtggta	gtgcacacct	gtaatcccag
24421	ctaactcagga	ggcggaggca	ggagaatggc	ttgaaccag	gagtcggagg	ttgcagtaag
24481	ccaagatcat	gccactacac	tccagcctgg	gtgacagaga	ctctgcctga	aaaaaaaaaa
24541	agaagcagta	tgccaggaaa	ctaaacacca	aataataact	tcacaatatc	gcacctacaa
24601	gatggtagat	attatgggtt	accattcaaa	aaataaaaaa	aagcttcaaa	ttgaaaacca
24661	ggcttatata	caattttgaa	aaagcaaaata	tctccccatg	aattttttatg	gggtgagtgt
24721	attagaattg	ctgtttcccc	tctgatatga	attaaatggc	tcccatttca	aatccatttt
24781	tgtgattgag	gcatactggg	aatgtttttt	ctgactctaa	atatagtttt	atttcagcga
24841	aggccttgaa	gtaactcctc	cagtcacaaa	gcatactgga	aatttgagtg	tgactcaagc
24901	tcattgtgaa	atgatgttta	ccaagttagt	attataagca	aagttagtta	ctctgttata
24961	aagagaatgt	tgcaaagggt	aacaagattt	taaagtatat	gttgagtatc	ctttatctga
25021	aatttttggg	gccaaaagtt	ttcagagttc	agatttggtt	gcatttttga	atatttgcac
25081	atacctaagg	agatatcctg	aagatgggag	ccaagtataa	acacaaaatg	cattatgatt
25141	catacatacc	ttgtacacat	aggctgcagg	taattttata	aaataacttta	aataatttta
25201	ttaaacatga	aacagagttt	gtgtaaagca	ctttggtatg	gaattctcca	cctctgggat
25261	catgtcaggg	ctcaaaaaaa	ttttagattt	tgtgattttg	gattatgatt	ggatgctcaa
25321	cctgtatata	tgtatatata	ggtatatata	tatatatata	tatatatata	tatatatata
25381	caccttcata	aaaatatttt	tcacaaatgt	aataataaat	aagctcYctc	aggttttctt
25441	ataaaaacatt	tttaaacgta	gctatcaaat	gcttctctgg	attaagtaac	attatctgaa
25501	atttggtttt	ctctactaca	ttattgctgg	taaattatga	ttttaagtac	aatgatgttg
25561	gccaaaatta	aggagcttta	taagaaactg	gcataatgaa	tttattttat	ttattctgga
25621	catactgttg	agggtaatcc	tgactccatt	aactccatta	aatagtcttt	tgttttttgt
25681	tttgttttgt	ttttgagacc	gggtggtgct	tgttgcccag	gctggagtgc	agtgacaaaa
25741	tcatagtctc	ttgtaacctt	ggagtccctg	gctcaagcta	tcctcctgcc	tcagcctctg
25801	ggacttacag	ccatatgcca	ctgcaactga	cttttttttt	tttttaattt	ttgtagacat
25861	ggggcMtggc	tatgttaccc	aggctggtct	tggactccta	gccttacaca	atcctcctgc
25921	tttaacctcc	caaagtgtct	ggattacaga	tttgaaccac	cacacctggc	ttagtcttgt
25981	attgttgaca	tccgccaaaa	ctcaccaatc	aattgatata	aaaaaaaaac	ctataatcac
26041	acaacataca	cagttttttt	atttagcttg	ctatggcaac	ggagaacata	cccagggaac
26101	cattaggagc	atcttagtaa	agtgaacta	gaaggggctt	cttatggact	ctggacaagt
26161	gttggagggc	tctgagaaca	gattatgaaa	gtaagtgtgg	tgctatcaga	gaagagaggc
26221	aattcgataa	tgggtatctc	aatgatgtcc	cccagaagca	gggacaatga	aaactaaaaa
26281	cacattgtca	ttgcaaggaa	gcagagatca	ctcagaaaag	gggatgtgag	tcattttatg
26341	gttgcagcac	agccttgctt	ttgttttagt	agaaacactg	ttgttactgg	aagcatcact
26401	aaatttcctc	tgaccatgtc	tgcatccatc	gaaaagagtg	gttggttttg	gttagaacca
26461	tactgatttg	caagcagggc	acatttttgt	ttcctcacga	atccaacaga	acagaaaatt
26521	attagtttct	actaccagag	Rtgcgtggtg	ttctcctcaa	ccactgggtt	gttttMatct
26581	tattaaggct	agaaattgta	aaaagagaga	aaagactgac	tccatcctct	aagtactgc
26641	tagaacctcc	caaggcaaca	cagctacagt	gcactctatt	gacaccaaYg	gaatcagtag
26701	catcatctta	attttcctaM	cttcttttcc	ctttacccaa	ttttatggct	tatttccatt
26761	ttttacttta	catgttttga	aatcacaccc	aatttttttt	taggctaaag	cagagagtaa
26821	ataaaaaata	actgtcaagc	ttggattgtc	ttttatttgt	gagctcagaa	tcaggatgag
26881	aagaaggtag	gaagttgtta	tgaattgtga	tggaaagtac	ggacaatgct	gggaaacaag
26941	ccagtattaa	ggactgtacc	cattttctaaa	ctgggctcca	tcgttaactg	cccttgtgac
27001	ctacagcaag	cattgaactt	tcttaagctt	tattttcctc	atgtttaaag	tagaataata
27061	ttgtaccaat	atagtttaga	taaaaataaga	attagcatgt	gaaagacttc	tcRtaatttt

```

27121  acatgacaat  gtatgaatat  tttgttaaaa  atttagtgcc  aaaYtctttg  tgcacaaaaa
27181  gcataYtggt  taaacaaact  ggtatctacc  tatcttagaa  atttttatta  ttaacagatt
27241  taaagccaaa  tataacaata  tacgcaaaag  taggagaatg  ggtttctaaa  tttttggaaa
27301  acaataatag  gagttcttaa  gatgcatgaa  aggcaactga  atttcatgac  agaaatagta
27361  agaaagaagg  aataagaaaa  tgtttcttgg  ccgggcgag  tggatcatgc  ctgtaatccc
27421  agctctttgg  gattacaagg  caggtgggtc  acctgtcaga  agttcaagac  cagcctagcc
27481  aatatgatga  aacctcgtct  ctactaaaaa  tacaaaaatt  agctgggcat  gatggccagt
27541  gcctgtaatc  ccagctactt  gggaggttga  ggcaggagaa  tcgcttgaa  ctgagaggtg
27601  gaggttgtag  tgagccgaga  tcgcgccact  gcactccagc  ctgagcgact  gagcgaatct
27661  ctgtctcaaa  aaaaaaaaaa  aaaaaaaaga  aaagaaaatg  tttctttcct  tgaaggtgac
27721  tataatgaaa  tctttcactt  ttctttcatc  gtagaagca  agaatcaca  ttatacagca
27781  attatgcaaa  acctcccatg  ctcataaata  ctaatcta  gagccctgat  tcagattcta
27841  aatcatccaa  ttaaattgcta  attggcttat  tctctttaa  atgagtttac  actgattcta
27901  ttcatttcct  gatctttgga  cgtatgagtc  tagtttcctt  gattaaaaag  gaccatggct
27961  ggatggtagt  aataagaaa  caagcacata  aagtatcaca  gagcccatg  gccaacatag
28021  taacataagt  tctatgaaga  cgaaaatat  actttacca  tatatgcctg  gattttacac
28081  gtcatactta  tcaaaattta  tatgcagttc  atctttgcaa  attgtcttca  gagacgggtt
28141  agaggcaaca  tacgataacc  atttttgcat  cagactctca  tctacatatc  aaccacacat
28201  agccttcttt  tccacagtca  cctctcattc  aattaaaatg  gctctgaatg  aatgaccttt
28261  tgggtgtttac  aaataccaag  Wctgcttga  agagcttaag  ctttgctgcc  attgaagaca
28321  agctccagac  agctccagaa  tgtgctatgg  ttgagttctg  cataagtctc  acctgatggg
28381  agtatcctca  gagaagagct  ggctgttgca  caggaggcaa  tgctccctgg  aaaggaaccc
28441  agctcctatg  cctgggaaac  agtcacattg  cttcagggtc  acgttttgaa  aatactgtgt
28501  atagcctatg  cataaaactc  ttgtaaaaca  tttgcaaaga  cacctgatac  tttctattct
28561  tagtcaaga  gttcaactca  ttttatttaa  aagatatata  aatattaaag  ttgcaagagg
28621  ctgggtattc  aaaaacctga  tattcaacag  aagtagagt  aattaagatg  gtgaagttt
28681  gcttccattc  acattctcca  tgaagccaag  agaccacaag  tctacctaga  agatttcgta
28741  tcttcccagg  taatcacttt  aaaagggtgac  tcagaccagt  gtgtaggtct  atgtttatct
28801  ttctgtccat  cctaaaccac  ccctgtgagc  aggatacaga  taaggaaaag  ggaagattgt
28861  ttcactagtc  tggcaggcaa  agtctagata  actgattgaa  agatgccatt  atgggacata
28921  ggctgtgttc  agcctaccta  gttgtaaggc  acacgataag  ataaaaatcc  caggaaagat
29041  acattttgga  aatcctgctt  ttgactcagt  tttcaacag  gcattgtcat  ggtgatgaga
29101  agagaaaaac  ttgggagaaa  acgctagaag  aaattaaaat  gatgtgacat  gatattctacR
29161  taagggtttc  cctgacatag  acctgtcag  gggccttgcc  agatatttat  gacactgtgg
29221  atgccatgtg  gttagtggaa  ctccagagt  tcctgtaagc  cacaccagca  tctgatctct
29281  gagtttctat  agggccatgg  ctagtagggt  gatagatctt  ggtgatactg  aaacgttaat
29341  acagccaagg  ccattttggg  tgtacatgta  ggtgtgtgcc  tatgtgggtg  aaattgctga
29401  ttttatctcc  cacaatatgc  tgagtgatca  tgaaagtaac  attttaaatg  tgagtccaga
29461  gtatcaactg  gttttatcaa  aacctaatg  gaaaaagtat  atatatatat  atatatatat
29521  atatatatat  atttatatat  attccacctc  tctaggatat  ttttgggtgt  ttttgggtgt
29581  ttttataatt  gttattttca  ttgtagtaag  agtattttaa  tgagatctac  ccttttaaca
29641  aattttaagt  gtacaataga  gtatcattga  tgaatataaa  tacattatta  aaaagatata
29701  tgttggtgag  gttatgtaga  aactggaacc  cttgtacatt  gttggtggga  atgtaacata
29761  gtgctgccac  tacggaaaat  agtatagagt  ttcctcaaaa  acttaaaaat  agaactgctg
29821  tatgatccag  caatcccact  gatYggtatt  tatctaaaag  aattgaaact  cagattttga
29881  agagatagta  tttgcacccc  tattttcatt  gcaacattat  ttacagcagc  caggatgtgg
29941  aaaagtgtcc  attgatggat  aaatgtgtaa  agaaaatgtg  gtatatagag  acaaggaatt
30001  ttattcagcc  tttaaaaaaa  ggaaattctg  caatgcaaca  acatgaagga  acacggagga
30061  cattatacta  agttaaataa  atgagacaca  gaagggccag  ttctaagctc  tcccgggctc
30121  gaaagcttaa  gtagatgaaa  aactcctccc  ttctcaggcc  cagccggaag  acgcaagacc
30181  acttgctgca  gcagcgtgag  tcagcaagg  agcagaagca  ggaagagagc  cagcaagaag
30241  acacctaccc  tggcagggaag  acacgtaccc  gtgaagatag  agaaagaggc  catccaggta
30301  caacgtagca  gttacctcag  aataggacac  ctccgttttc  caggagacta  taaaaccttt
30361  gccccctcct  cacttggggg  ctgacggcat  taggcctcag  cccgcctgca  cccaggcgct
30421  cactaaaaca  gcatgttgcc  cggcgcggtg  gttcatgcct  gtaatcccag  cactttggga
30481  ggcctaggcg  ggcggatctc  gaggtcagga  gatctagacc  atcctggcta  acagggggaa
30541  accctgtctc  tactaacaat  acaaaaaatt  agccgggcat  gatggcgagg  gcctgtagtc
30601  ccagctactg  gggaagctga  ggcaggagaa  tggcgcgaa  ccgggaggcg  gagcttgca
30661  tgagKggaga  tggcgctact  gcaactccag  ctggcgggca  gagcgagact  ccgtctcaa
30721  acaaaaaaaa  aacagcatgt  tgcctccacac  tgcctgtgt  tgcctgtgtg  cgcactcttg
30781  gtgttcgaac  tgatacaaga  accttacacc  aatactgtat  gattcttttR  tatgatgaat
30841  gtaaaacagt  aaaactctta  aatgcagaga  gtagaatgg  agttactagg  gactaaaaag
30901  agggggtaac  tgggaattac  tattcagcag  gtacaaaatt  tcagttaacc  aaaatgaatt

```

30961	agttctagag	atctgctgca	caacactatg	cctattgcca	ttgtcgtttt	cttgttttgg
31021	tttgggtttg	ttttatgctt	aatttcttct	taaaaaaaa	Kattgactca	agccctattt
31081	agctgtccag	gaaatctatg	ggggaacaaa	gaactataac	acaccagtca	cccagtctgt
31141	cccctttcca	ctctgttaga	gtgagatgtc	ttagatagat	ttgatctcag	gctccaatcc
31201	tgctactcag	ttatgcttag	cattccaaga	tccttagcaa	gacataaacg	tcctgtcact
31261	gtggctgcta	ccaaccactc	tagctgcaac	tcttgctctc	tccttatatg	cagtctaggt
31321	tccaacagta	ctgaaccact	gaattccttg	aatttgaata	cgatttaaca	cttccttgcc
31381	tcatgtgaca	ttctttggtc	agaatccatt	tcctggctct	cttttttgcc	ttgataatag
31441	tagctgtatt	cctaggctgc	ttatttacat	atatcttctg	ggaagccttt	ctttttaaaa
31501	tttgtacaaa	tttatggggt	atRgtgact	ttttttatgt	gtttatagca	tggagtgtat
31561	tcaagtcagg	gtattcaggg	tgtccattac	cccagtacaa	tacgtttttg	ttaaactatag
31621	tcactctgct	atcaaacatt	ggatttattc	cttctattta	actgtatgtt	tgtaccattt
31681	aaccagtcct	tttcatcttt	ccccgtctcc	ctgactcacc	tttcccatte	tctgttgtct
31741	gtctttactc	tctactttta	ggtgacaaa	cttttttagct	cccacatatc	agtgagaaaa
31801	tacaataatt	gcctttttgt	atctggctta	tttcatttaa	gataatgact	ccagtttgat
31861	gtttatgcaa	atgacattac	ttcattctat	tttatggctg	aatagttttc	cattttgttt
31921	atgtatgcc	cattttttgtg	agccattcat	cattgatgga	catttaggtt	gattttatat
31981	ctttgctatt	ctgagtagtg	ctgcaataaa	catgttagtg	caggtatctc	tttgatatac
32041	agattttctt	ctgtgaagcc	ttttgtgatt	taacctccag	tttcatctat	tagacatcta
32101	ttatttactg	accagtatct	attctaacaa	cattatcatg	gacatcctta	aataaatgtc
32161	tctactaca	ggtttcaagt	ttgctggga	ttgaagggtt	ggaggatgtg	ggacttccaa
32221	tgcaaaaact	gggacagtga	aaaccaggat	agttggtcac	cctaccatgg	atgggacctg
32281	accttctgtc	tagtcacttt	tgtttaaaat	atatttcccc	cattaRgctt	tgatgttctt
32341	tgaagRcaat	gtcctccYgt	tggctccttg	catataatac	ttttcttgat	gcttaataaaa
32401	tattcagtg	atgattttct	attttcaaa	ggcttaaaat	agctcttggt	gtatgaacat
32461	atttatttta	tgggttgttt	ccttctgttt	aatagtagca	ctgtgatatg	gtttgtctgt
32521	gtccccaccc	aaatcttata	ttgaattccc	atgtctttcc	cttgctgttc	tcatgataga
32581	gagtaagtct	cacaagatct	gatggttatt	ataaggggga	attttctctg	acaagctctc
32641	tttttgtgtg	ctgccatcca	tgtaagacat	gactggctgc	tccttgctct	ccaccgtgat
32701	tgtgaagcct	caccagccgt	gtggaactgt	aagtccaata	aacttctttc	ttttgtaaat
32761	tgcccggctc	caggtgtgtc	tttatcagca	gcatgaaaat	ggactattac	agtaaatgtg
32821	tacaagtaga	gtgggtgtgt	gctgaaaaga	tacgtgaaaa	tgtggaagca	actttggaac
32881	tgggttatag	gcagaagtgt	gaacagtttg	gagggctcag	aagaagacag	gaaaatgtgg
32941	gaaagtttgg	aaccgtagag	acttgttgaa	tggctctgat	caaaatcctg	atagccatat
33001	ggacaataag	gtccaggctg	aggtgggtcta	agatggagat	gaggaacttc	ttgggaactg
33061	gagcgaaggt	gactcttgct	atgttttagc	aaacacactg	gtggcatttt	tcccctgctt
33121	tagagatttg	tggaaacttg	aacatgaaag	agataattta	gcgtatctgg	tgaagaaat
33181	ttctaaacag	caaagcattc	aagaggcgac	ttgggtcctg	ttaaaggcat	tcaattttat
33241	aggggaagcag	agcataaaa	ttcagaaaat	ttgaagcctg	acaatgtgat	agaaaacaaa
33301	agccccattt	ctgaggagaa	attgaagctg	gctgcagata	tttgcataag	taatgaggag
33361	cctaattgttc	atccccaaga	cagttaaagaa	aatgtctcca	gggcatgtca	gaggtcttct
33421	tgtcagcccc	tcccatcaca	gactaggagg	cctaggagaa	aatggtttca	tggggccaggc
33481	ccagggtttcc	tgtgctgtgt	gcaatctatg	aagttggttg	cctgcatcct	ggccactcca
33541	gccatgacta	acagggggcca	aggtacagct	caagctgttg	cttcagagga	tgggaagccc
33601	agcccttggc	agcttacatg	tgggtgtgag	cctgtgggtg	cccagaaatc	aagaattgag
33661	gttttggaac	ctctgtctag	atttcagaag	agttatagaa	atgcctggat	gcccagacaa
33721	aagtttgctg	cagggatggg	gccctcatgg	agaacctctg	ctagggcagt	gtggaaggga
33781	aatgtggggg	tggagtccca	acacagagtc	gctactgggg	cactgcctag	cagagcttga
33841	gaagagggcc	accatcctcc	agacctcaga	atggtatagc	caccgacagc	ttgactgtgt
33901	cacctggaaa	agctacagac	actcaacacc	agcccgtaaa	agcagctggg	agggaggctg
33961	taccttgcaa	agccacaggg	gtggagctgc	ccaagaccaa	ggcagcacac	ttcttgcatc
34021	agcgtgacct	gcgtgtgaca	catggagtca	aaggagatca	tttgagtttt	taagatttga
34081	ctgccccggg	ccgggcgcg	tggctcacgc	ctgtaatccc	agcactttgg	gaggccgagg
34141	cgggtggatc	atgaggtcag	gagatcgaga	ccatcctggc	taacaagggt	aaaccccgtc
34201	tctactaaaa	atacaaaaaa	ttagccgggg	gcggtggcgg	gcgcctgtag	tcccagctac
34261	tccaggagct	gagggcagg	aatggcgtga	accggggagg	cggagcttgc	agtgacttga
34321	gatttgcgcca	ctgcagtcgg	cagtcggggc	tgggcgacag	agcgagattc	cgtctcaaaa
34381	aaaaaaaaaa	aaaagatttg	actgccccac	aggatttcag	acttgcatgg	ggcctgtagc
34441	ccttttgttt	tggccaatgt	ctcccatttt	ggatggctgt	atttacccaa	tggctgtacc
34501	cccatttgtat	ctaggaggta	actaacttgc	ttttgatttt	ataggctcat	aggtggaagt
34561	gatttgcctt	gtcttggatg	agatttttga	ctgtggactt	ttgagttaat	gctgaaatga
34621	gttaagacct	tggggaactg	ttgagaaggc	atgattgggt	ttgaaatgtg	agaacatgat
34681	atttgagagg	gaccaggctg	gaatgatatg	gtttgtctgt	gtctccactc	aaatctcatc
34741	ttgaattccc	atgtttttgtg	ggagggactt	ggagggaggt	aatttaataca	tgggggcaga

34801	tctttcccat	gctgttcttg	tgatagaaag	taagtctcat	aggatccgat	ggttactgta
34861	agggggagtt	ttcctgcaca	agctctcttt	ttgcctgctg	ctatccacRt	aagacatgac
34921	tggtctctcc	ttgccttctg	ccatgattgt	gaagcctcct	ctgccatgtg	gaacYgtaag
34981	tccaataaac	ttctttcttt	tgtaaatttc	ccagtccttg	gtatgtcttt	atcagcagca
35041	tgaaaactga	ctaatacaca	ctgttRataa	gagtttaatt	aaaaaccccc	agaaaMagtc
35101	taaatgtcca	ttaaccaa	aatgtataag	tacatgttga	ataaatatac	tatggtccag
35161	tcactggaat	agtatacaga	aacaaaaatt	agttgactac	caatgtacaa	agcagattga
35221	tgaaatcttg	aattataatt	tgtattatat	ttggaagata	caaaagaata	cataataaaa
35281	tacgtatcaa	atatatttta	atatctgcta	aaggggaaaa	gaaaaaagac	ttccattatc
35341	aatcaatatg	tgtgggattt	tattattttc	atcaRtacta	ttacttttag	aagcagctac
35401	tactattgaa	cattcattcc	atgccacaca	tttagccctt	aatatacatt	atcttatttg
35461	attttacttt	ttcaataaact	ctgcactcta	taggtgaaga	aaccaagata	ctgaggtagt
35521	aaatgcctta	ccaagatatt	gaggtagtaa	atgccttacc	ttgccttgcc	aaaccaagat
35581	actgaggtaa	taaatgccat	ataaatatgc	tgttagtgtt	aaagccagat	ttcaatgctg
35641	gtcctacctg	gattcaaaac	taacatttgt	tcaatgtact	atattaataa	tccatcta
35701	gatctcaatt	atctgtta	gtgaatgctc	tttgtgtaa	acacactgca	ggggtccccc
35761	tcctcctatc	tcctgatcc	ttgcctccca	gtatttatgt	ctttgtataa	tctttctg
35821	ttgagttgag	gcaggacctg	ttacttgctt	ctaaccagta	gaatatggca	aaggtgatag
35881	atgtcacttc	tgtgattaca	ttacattata	taggattcca	tcttgctatc	agaattattc
35941	tgagagactgt	ctccttaggc	ccatatgaca	gccctcagct	caagatagtc	tccagcta
36001	ttccagcaag	aagccagggc	tctgagtcct	agaggcacat	ggaaatgaat	tctgccaa
36061	gtcagcttga	aagaggatgc	ttctctagtt	gagcctccac	atgagaaccc	agcccagctg
36121	acggcttgag	tacagcttta	taagacccta	agcagaggac	ccccgcttaa	tttgtgcctt
36181	aattcctgcc	ctaaaaacaca	attaggtaat	aaatgggtgc	tgttctcagt	tgctatgctt
36241	gtgatagttt	gttatgctac	aaaagaaaat	gaatacacta	tcttttttaa	aaaatcattt
36301	tttYcctcaa	agacaagtta	tgcacttatt	atgggagatt	attctacatg	tttatagaag
36361	gccttaacca	gagtttgga	ctggaactgc	tgccctgatga	tgtaattata	ggaattatat
36421	cccaaaaatt	agtaatat	tgaacttgt	aagagtacaa	aataatatct	ttgtgcatac
36481	tgcccacttt	gaagctcagt	tagtgctact	ttgactagg	atgggtgcat	gctgagggtc
36541	aggttccagc	cccagctgag	ggttgagggg	agtgggtgga	catggggcag	ggagctggaa
36601	gaacacttga	gagacagcag	gtaaatgaga	catggcttta	ttcagaagcc	tcttctcagg
36661	gtcagtggtta	catttatata	ctacacaaac	aatagtggct	gagagccaga	tWgggtgctt
36721	ttctatgtta	tgtcaacatg	gctatgatta	tataaggcat	gggactatgY	gcctgtgcta
36781	caaactcgct	gagtcattcca	ggctgtttac	cttgccctat	gcctgctgcc	ctatgcctgc
36841	ttggctgcag	cacagccatg	ttccttacat	atggcattgt	tgaaattgaa	atgtcatage
36901	cacactgcag	aagagagaca	tcaccagtc	ctcattagac	aacatagctg	tattaaatcc
36961	tgctctcttc	agaagccaag	acctatgcat	tccttcatga	atagctttta	aagggttaatt
37021	aaaagcagtg	agtaataaaa	atttataagc	ttctgcttat	aattttta	ggtaccattt
37081	ctgacctttt	gatttgtgatc	ccatgatgta	ctatccagat	gccagttgtt	cactgttccc
37141	aacaacttaa	aggagaactc	aacagaacat	agagtccttg	cttttccagg	aaagtttctt
37201	caaggtttgt	tctggtgatg	acaaattcct	atggctgtta	ggttttctgt	tgtgaatggg
37261	atattctttt	cttgagatt	cacaatctac	attagcattt	aataaatatc	aaaaagtttc
37321	aactaagaaa	aataattaac	aRtttcttta	ctagcattac	atgtacttat	ttgtcctcag
37381	aaagctttct	caaaaacacc	ttaaaactt	aRtaactcct	gcttgagaaa	gagtggtgta
37441	aaagaatgaa	ctcatcaatt	tggaatcttt	attttaaaaa	atatttacag	tagccttcac
37501	acttgaaatt	tgaagcatca	aatctattaa	ataacaaaac	aatatcagag	caagtaaaac
37561	taaaagttca	catggaaaaa	tttaacaata	aaacattatt	agaaactaag	tagtaatatg
37621	tgaccaagac	tggtcaagag	ataaacttgc	acatattaat	gctgttatgc	tcagctttta
37681	tgtcccagat	aagactaata	aagaatgatg	tcaatgttag	gaaggagat	taactacata
37741	gcaatatatt	tgtattgtca	atttgaatat	tttactgaaa	tatcagcaaa	tggatatata
37801	ttgaattttc	taacacagaa	gtaaattgtc	ttctgtcttc	aaattttatt	tgtattcacc
37861	aacaccaata	aggcaggggt	atgYtcccac	atgccaaact	agctacagtg	atctctttgc
37921	agtttcataa	acacagcacc	atctcccacc	ttagcactta	ctgtctttct	gcctgaaatt
37981	ctcttcctcc	aaatgtctga	atggcttact	ccctgtttt	atccaggtat	ttgtcctaat
38041	atctcatcag	ctctttcctg	gctaccttcc	aaaaagKggc	acatcaccag	cctcacttcc
38101	tcactccctt	actctttata	ttttctttaa	ttgaaataat	ctatacttga	gatagtttct
38161	gttggtcttac	ttgcttattg	actattttacc	ctagataccc	aactgcaata	taggtttctt
38221	gacagcagag	attttggttca	gttggtgatc	cccagcatgt	agcatagagc	ataattctca
38281	ggaagtatga	aataaatagY	Rcatgaatta	atggaatagc	agaagtcat	ccttcagaag
38341	agacttatgg	gcagtcgttt	aaggtaagtt	taacatagga	aaagcagtga	aattcaccat
38401	attaagacta	aaaaagaaat	atgcaatcat	ctcaatgcat	ataaaaaaac	atttaacaaa
38461	atctgacata	tattttta	gaaaactgtg	cacacactag	aaatagaagg	gaaccctctc
38521	aacctgataa	tcataactac	aagaaaagct	acagctaaca	tcacaatggt	aaaagactag
38581	atgctagatg	ctttcttcta	atatttgaaa	caaggcaagg	atgtccactg	ccagcactta

38641	tattcaacat	tacactagag	gtcctagtca	gtaaaataag	gtgagaaaaa	aattaataaa
38701	agccctccag	aaaggaaaaa	gagcaaatga	aacaagtcta	tttacagatg	acataattgt
38761	ctatatagag	catttatatgg	agtctattaa	aatgttacta	gatccaatat	gtgaatttag
38821	caagatggca	gagtaagaaa	tgaacacatac	aaaatttatt	gtgtttctat	aaactagtga
38881	tgaataattg	gaaaaatgaaa	ttttaaaaca	ttatcattgc	taccatgagc	caataaattg
38941	aaaaatttaa	gataaatgga	taaattccta	gatactatca	agattgaacc	aggaagaaat
39001	tcaaaacctg	aacagaccaa	aaacaagcaa	tgagttcaaa	gccataataa	aaagtctact
39061	agtagaggct	gggtgcggtg	gctgatgcct	gtaatcccag	cactttggga	ggctgagggtg
39121	ggcagatcac	aagggtcagga	gatcgagacc	atcctggcta	acacagtga	accctgtctc
39181	tactaaaaat	acaaaaaaaa	aaaaaaatta	ggcatggtgg	caggtgcctg	tagtcccagc
39241	tactcgggag	gctgagtcag	gagaatggca	tgaaccagg	agacggagct	tgcatgagc
39301	tgagatcgtg	ccactgcact	ccagcctgag	tgacagagtg	agactccatc	tcaaaaaata
39361	aaaaataaaa	aaagtctact	agtagagaaa	agcccaggac	ctgatgactt	cactgctgaa
39421	ttccaccaaa	tatttaaaga	agaactaata	caaatcctat	tccaactatc	atgaaaaata
39481	gaggaagagg	gaatacttcc	aaactcattt	tactaagaca	gtattaccct	aataatcaaaa
39541	ccagaaaaag	acaaatcaaa	accagaaaaa	aaaatcccaa	aaatcctaaa	atttatatgg
39601	aaccacaaaa	gacccagaat	agccaaagct	atcttaagta	aaaataagaa	aacaggagga
39661	agcatattaa	ctgccttcaa	attatactac	agagctgtag	taaccaaacc	agcatggtag
39721	tggcaaaaac	agacatagac	taatgaccca	gaatagagag	cctacaaata	aatctacaca
39781	tctatggtga	acttggtttt	gacaaagatg	ccaagaacat	acactgggga	aaagagtctc
39841	ttcaatcaat	ggtgctggga	aaactggata	ttcataagca	gaagaataaa	actaatggcc
39901	tatctctcac	catatacaaa	aaaataaaaat	ccaaatgaat	taaagactta	aatctaagtc
39961	atcagactga	aactactaca	agaaaacatt	gggcaaaatc	cccaggcatt	ggtctgggca
40021	aagatttctt	gaaaaatacc	ccacaagtac	aggcaaccaa	agaaaacatg	gacaaatggg
40081	atcacatcaa	gttaaaaagc	ttctggacag	caaaggatac	aatcaacaaa	gtgaggagac
40141	aaccacaga	atggaagtaa	atatttgcaa	actaccctc	tgacaaggga	ctaattgcc
40201	aaatacatga	ggaactcaga	caactccaca	ggaaaaaaat	ctaataatcc	aatcaaaaat
40261	gggcaaaaga	ttttaatagt	tatttctcgg	aggaaaacat	acaaatatca	aacagacatt
40321	caaaagggtg	tcaacatcat	tgatcatcag	agaaatgcaa	atcaaaaacta	caatgagata
40381	tcatctcact	ccagttaaaa	tgttatagac	aggcaataac	aaatgctgga	aaggattttg
40441	ctctgtacgc	gtgtggcggg	aatgttaaatt	agtacaacta	agatgaagaa	gagtgtggag
40501	gttctcctcaa	aaactgaaaa	ttgaactacc	atatgatcca	tcaatccaac	ttctgggtat
40561	acacctaaaa	gaagcaaaaa	ttgtatatgt	aagatatatc	tgcaactcta	ttttgttg
40621	agccctattg	attataggca	agatttagaa	gcaacctaaa	tgttcatcaa	cggatgaatg
40681	gaaaaataaa	atgtagtaca	tatacacaat	ggcgactact	ttaaccatta	aaaaaaatga
40741	gatcctgtca	tttgcaacaa	catgaatgga	actggaatga	atgttaagtg	aaataataaa
40801	taagccaggc	ataaaaaggc	caacatcaca	tgttctcact	tatttgtaga	atctaaaaat
40861	caaatcaatt	gatctcatgg	acatagagag	ttgaaagatg	gttaccagag	gctgggaagg
40921	ggagtaggg	gagggcaagg	gtggggatgg	ttaatgggta	ctaaaaataa	tagaaaaaaa
40981	tgaataagac	aaactaatag	cacaatcagg	taattatagt	caataataat	tgtatatatt
41041	aaaataactt	aatataattg	gattgttttt	aaactcaagg	ataaatgctt	gaggggatgg
41101	atacccccatt	cttcatgatg	tgcttatttc	acgttgcatg	cctgtatcaa	aacacctcat
41161	gtacctcatg	gatattttaca	cctactacRt	acccaaaaaa	ttctaaaaat	aacttttttt
41221	aaatattgtc	atttataaca	gcaccaaatt	tatcaaatat	ttagaaacat	ttttgataac
41281	attttaaagt	gtgtacactc	aaaactgcaa	aaattgctga	gagaaaaatt	taaaaatacc
41341	taaatagagt	gatactgtgt	taatggattg	gaagagctaa	aaattattaa	gatgttggtt
41401	ctcttcaaac	tgacttttta	attttttttt	attatacttt	aagttctagg	gtacatgtgc
41461	acaatgtgca	ggtttggttac	atatgtacat	gttggtgtgc	tgaccccata	aactcgtcat
41521	ttacattagg	tatatctcct	aatgctatgc	ctccccctc	ccccacccc	acaacaggcc
41581	cgggtgtgtg	atgttcccc	tcctgtgtcc	aagtgttctc	tttgttcaat	tcccacctat
41641	cagtgagaac	atgtgggtgt	tggttttttag	attttaattaa	atcccaataa	aaataatagc
41701	aggctttact	gtagaacttg	gcaagctgat	tcaaagttta	catgtaaaca	aaggacctag
41761	aatagtcaag	atagctttta	aaagaggaaa	gtctgaacac	atgcaacacc	taagtttcaa
41821	gacatattat	gaagctataa	taatcgagag	agagtggtaa	tggcataaaa	gcagatgtat
41881	agatcaataa	aagagaaaaa	aaagtccaaa	aatagagttt	cacttacatg	gtcaagatat
41941	tttcacaaaag	gtaccaagg	aattcgatag	ggaaaataat	ggcttttcaa	caaagtgttc
42001	taaaattggt	agatattttac	atgtaaaaaa	aatgagacct	gattcatata	ttgtaccctt
42061	ataaaaaatta	actcaaaatg	tatcacagaa	taaatacaaa	actaaaaata	aagacttgta
42121	ggagaaaaata	tttatgaccc	aggtttaagc	agagatttct	tagatgtaac	ttgaagagaa
42181	taatccataa	aattaaaaaca	tgaaaaaaat	agttgggctt	aatctatgtt	aaacatttct
42241	ctctcttggg	agacattgat	aagagaataa	acaaacaagt	caaagactgg	aataaatatt
42301	tgaaaatcgc	atattggaat	tgtaaaccaat	taagggaattg	tgaccaaacc	atacaaaaaat
42361	ctctcaaaac	acaacaaaat	gaacacagag	aactcaatct	aaaaactggg	taagaaattt
42421	aaacagttac	ttcactagag	aaaatacatg	gatggaaaaa	aatgaataa	aaagatactc

42481	aatattat	ataattagg	aaataaaat	aaaaccacaa	ttagatacca	ctat
42541	atcagaatgc	ttaaaattaa	aaagattaat	caaagtgttg	gcaaagctgt	ggagcaactg
42601	gaactttcac	acagtgtat	ttgtaatgaa	atatggtaca	gccactttgg	aaaacagttt
42661	attaat	tcaaaagtta	aacatatatc	tttcctatga	ctcagctatt	ccactgggca
42721	tttatgcaag	agagataaaa	gcata	aaacaaat	ttgtatatga	atgttcata
42781	cagctttatt	ttaaacagct	gaaacctaga	agcaatccaa	atgttcttca	acaggtgagt
42841	ggattaaaaa	aatgtggaat	at	tagaatgtt	ttcaaccata	taaaagaagg
42901	aactgttgac	acatctaatt	acctgaataa	atatcaaatt	at	aagtgaacca
42961	ggtaaaaaag	aaggcatatg	at	tttataca	tattctagaa	aatgcaaaact
43021	aatctattgt	gacagaatgt	ggtttgtatt	tgcatgatgt	aagacaagaa	aggggtgggga
43081	aagggatggc	aaagaatcac	aaggacactt	ttagaagtaa	taaatacact	tactat
43141	atcatagtgg	tagtttcaca	ggtgtgtgac	atatgccaaa	acatcaaatt	gagcacttga
43201	aatatatgcc	at	tatctgttac	acttccacaa	gcctttaaaa	gaggaagaga
43261	gagagcaaat	ttgaagtgtg	agagatagaa	tgaaattgct	ttctgcttgt	ttctaccatg
43321	atgccctcat	tcaataagct	ggttacattt	cctcatgggt	ttgtaggaca	tataatgaga
43381	atatatatat	atatatatat	acacgtata	tattatatat	agactatgta	tgtgttatat
43441	atacatatat	gtat	atatatacat	ttaatcgtat	acataataac	at
43501	acttctttaa	aaaggcctct	gagagacccc	ttgttcttct	caccatgtga	gggtatagca
43561	agaagggtgac	atctgtgagg	aagtgggccc	tcaccagtca	ccaaatttgt	caacaccttg
43621	atcttggatt	tctcagcctc	cagaactgag	agagtgaagt	agagagagag	agagagagag
43681	agagaagtgg	tagcacatg	cctcattttg	attactttgt	tgactgattc	aaaatcatga
43741	agttgaatct	ttaaactatt	cgtagcctga	aatggtttgt	ggttttagaa	tttcttaacc
43801	tcaactatcc	cttctcaga	aaaactcact	tggaaatttg	tttaaattat	taaaagtaat
43861	cccaaaggag	agtggcctag	tcaggtcaat	catgattccc	gtttctaagg	caatgcatct
43921	agattcaaaa	aataggagta	gatatttcca	acttacagtt	atcttaatca	ttggactgtg
43981	cccagagtct	atcatagtgt	ctgttgcctc	cagcatatat	ttcactatgg	atcgaactcc
44041	aattatgctt	tctataat	tctttaagtgt	ctaacaagca	atgatactca	aatatgtata
44101	ttgagattct	cattttcaac	tgctttctga	acatatgtgc	caagatatct	tctgaaattg
44161	aagatgcttt	aaaaacattt	atgacattcc	ccttccaatt	attat	ttattattga
44221	ggcagagtct	cactctgtca	tccaggctaa	agtgcagcgg	tgcgatctcg	gctcactgca
44281	acttcacact	cctgggctca	agcgtatctc	ctgcctcagt	ctcctgggta	ctcgggatta
44341	cagtcatgcg	ccaccaaaacc	tgactaat	ttgtat	agtagagacg	gggttttgcc
44401	atgttgaYca	ggctggctct	aaactcctga	cctgaagtga	tccgcctgcc	tctcttgcaa
44461	tattttattc	ccgtcttact	ttcctat	ttMgtgatca	ttaaaaatcc	agaagcatat
44521	attagggtaca	gtgtggcaaa	aaatagaaat	gatcaaatca	ttttgagggc	Rtcaatgaaa
44581	gctttcagtg	aagttgggtg	aactgcgcct	tgcaagaa	ggaagaagaa	ttctctaagt
44641	aaacagaaag	ttgcattgct	atggtctgcr	tgcttggtgc	tccccaaaat	atacaaaata
44701	tatatgttaa	aatcctgcct	ccaaagggtga	ttgtattaga	aagtgggact	tttgggaggt
44761	at	taaaggcaga	ttcctaata	atgggattag	cacatttata	aaaagggtccc
44821	tgagaaaccc	tttgtccctt	ccactatgtg	agtacacagc	aaaaagggtg	catctgtgaa
44881	gaagtgggct	ctcaccagac	accaaat	ccagacctt	gatcttgagc	ttccagcct
44941	ccagaactct	taaatataaa	tttgtgtgtg	ttataaaacta	tccagggttat	gggtat
45001	tatagcagct	cagatgaact	aagacatgca	ttcaaaaaca	aagaaaactg	cataaaccaa
45061	tgcatgaact	agaagcagtg	gcattgtgct	tggggatctg	caagaagttc	tatatggcta
45121	ttgcacagca	tagagactgt	atgtagaagg	aaggtacaca	gacaggaaga	gggtcccaga
45181	gaggagtaat	ggatat	tgtaatttaa	aataattcRg	tacagacaga	gggatatatg
45241	tgtttatctta	caacttaYta	taaactctaca	catctacacc	cttgagtaat	tggttaacat
45301	ttgagaggaa	ttttccta	attgcagttc	cccaggaaac	actcctgagt	cctaagattc
45361	catgtaggag	gttactgtgg	agcgttctca	agaaccacag	gtgagagaag	ggaagaaagc
45421	cctatctgat	gcagtggtgaa	cagataagtc	agctaaggaa	gagagctccg	gggttggaat
45481	gactcttcag	agttatgtta	attaaagcaa	gtgggctgga	ccagaccttt	gtgctgtcca
45541	catcaagcaa	tcattggctt	ccagctgcct	ccaagaggag	gctggtcttt	tgtctgttta
45601	tattattgag	ctatgctgca	gcttgggttc	tccagaagac	aaaaactgaa	acagagttag
45661	gagtgcaggt	taattttgccc	atgtggcaga	cactgagacc	tacttaggag	tgcacacatc
45721	cattagggaa	tcaaactgtg	gagagaaatg	aggcagaagc	tggcttaggc	aaggggagga
45781	acagaccacc	atgtggagct	gacaaagtct	ctcttagccc	agtaggga	tccagaccaa
45841	agattattga	ttacaggagt	cccgtgatgg	gtggaaatta	ctaggccatt	gtagcaccac
45901	tttgttcagt	ctctggccca	aggccaccat	gagaagagtg	ttaccttggt	tcaaaaatgg
45961	aggggaagagc	aaactaaaaa	agccaaaagc	tggaggatgt	cagcgaacca	cactcctcgc
46021	atctagacag	taagggtctt	cttgaaggaa	gatctgagct	gcacacctcc	tatgtttacc
46081	atgtctgcct	catgacttca	tgatctctct	cttctctctc	cttttttttt	tttttgagac
46141	ggagtttcac	tcttgtcacc	caggctggag	tgcat	caatctcagc	tactgcaac
46201	ctccacctcc	cagattcaat	cgattctcct	gcctcaggtc	cccagtagc	tgggatgcct
46261	cgtgatcgtc	taagatctat	ggtgtagtta	tgggtcctc	cccttttgtg	gaatctctct

46321	gtccgggttcg	tgagaacctc	caaaaatata	tggtatttga	ctaaccattc	cttgaagttt
46381	cttccagggc	tgggatttta	ttctttttta	gtttccctga	gatattgcaa	atacaaatgg
46441	caaatgaaca	tataatgaag	tcactatttt	agatgtacaa	ataccacaca	cactgggtacc
46501	tctgagggtta	tttgaactgc	tgaatagagt	aactttttatg	tttccaattc	ttgctaattgg
46561	agagagaaga	ttgcctctgt	gttaatgttg	gatttttaaaa	tatatccttt	ttaagcacta
46621	ttaagcatgt	tttaggttac	aaagtataat	ttattcatct	tttccattat	aacagaaaaa
46681	atattataag	gaaatgtctc	aaacttttct	ttcttacagt	tcagtctcat	tgaaaagaaa
46741	gatgtctaca	tcattgagaa	gaaaaagaag	cccaatgtgg	attcagcaca	attaacacgg
46801	agtacttgct	aaggcacaga	cactgcagga	ataaattgta	ggctgcattt	ctgggtcttag
46861	gactcagtc	aatgggacat	acaagccatg	caagccaaca	gcataaattga	agacaaggtg
46921	ttggaaaaac	aatatgaaca	agtttttata	ggatcaaaag	aaagaaagaa	ttcgaattga
46981	tctgaggagt	acaagaatat	tccagaattg	gggaataaat	aggagagatc	aacagaaagg
47041	ccatttgaat	cagaaccagc	aggcaggaag	tatgaaacca	tctagcagcc	tgttctgtaa
47101	ggacaaaaga	tcattaaata	ggcaaattgc	agaaagccag	ggatagtagt	gttttagcttt
47161	ctttaattcc	tctttgttga	gtttcccat	gtcttagaaa	ttattcagga	ggtaagacct
47221	actggtgtat	ggcacagcag	gaatatacag	cctattaatc	Rcaatcaagt	ctctaattgct
47281	tgtatttaata	aatcatgttt	aatgtagaat	aatttctgaa	aataatgcca	cagatcaata
47341	aataaatctc	taagaaaagc	tgaagtccc	agggtcatat	ctgagaactt	gaggccaggg
47401	gagtttagac	actctaagta	gaagacaccc	aggcaagact	tttgtggata	aggacaatta
47461	tacaagttca	tagccttctg	aaataagctt	agttctaata	atctcaggaa	aaacacaaaa
47521	taatttagct	atactgatgg	tttttgagga	acaattttcc	ccagatttaa	aaagattgat
47581	ggatgtggct	aaatcactgt	ccctaaccct	tttatagaac	ccagaggttc	cagccttaag
47641	atttaatttc	cagggttagg	atgcaagaag	tgacttactt	ttgagttcta	ccaggaagca
47701	aaatggcaca	catcaaaaat	aactcgtttg	gcgatttcat	gagtttccaa	gcactttata
47761	aaatgtacag	tcatacagct	ttttcttatt	gcaataagaa	actattttaa	tgctagaaaa
47821	ttgtatatatt	ccagcttcaa	aatgaataat	gtcatccaac	agaattaaag	cagtaactac
47881	agaaagagta	atcaataaat	aggctcttcc	tcttactacc	ctagagaatg	gggagatgtt
47941	atttgatcaa	ttgtttttct	aaaatttatg	actgttagaa	gccaagtcta	aatttatata
48001	gaaaaagttc	aaggaaaatga	tgtaactatt	ctctaagaag	gtgattttaa	cctctggatt
48061	acaggagggt	taagatgttg	cccttggtgt	aatgcagtg	tttaacttaa	gggccatcgt
48121	gtaccatcaat	tttccctcat	gaacacaaatc	accagccaac	caaaaacctct	actatttcta
48181	tatttctctga	aggactcttc	cattcccaaa	ttgagtgtat	gaattgcagc	ctcataaagg
48241	taccaccgaa	atgtggcagt	ctcatttgga	gacacacatt	acaaaagagt	ttcagccaaa
48301	atagtttgaa	tcagaaaagt	gtccttgaat	agtgattcag	cctaatacat	cactaagaag
48361	gagaaacaaa	tgaggcattg	gagaccttcc	atccaagtgt	caaataaggaa	aagtcgttca
48421	tgtaccatgg	catgtacagg	aaagcagtga	tgtttcaaga	actaactggc	aaMgactatc
48481	agattttaga	aacaaagtct	gcatttcatg	atgccccaaa	atctagttga	aaacataata
48541	ttcaatatatt	gttacctctt	tttttgata	tttatttggt	atactgaact	gaattattca
48601	cttattattt	gaagttatgt	tcttatagga	agctaggcac	caaccagcag	ttgctaataa
48661	tatactataa	tttagtaagt	aattcaattg	aataaaattt	atgacaaatt	taacaaacat
48721	gttaattaaa	atgcatactc	acgtctgcagc	gtcacatttaa	tctttgtgcc	Rccagtgctt
48781	atgccatgct	tagtatgcat	caaataattg	agcagtacac	aagtgagtac	tctgagagct
48841	ccccccacca	aaaatatgat	gattaaatac	agttatgac	agatccccag	agtggtggctc
48901	taaactgtat	ggggggccaa	tttgaatact	gttggtgtct	acactgttat	tacctatcca
48961	gtatctattt	ccccatattc	cttataaata	aaacctagat	tttgattggg	acagtaaggt
49021	gtcccactga	aaactcattt	ctctaaccaa	tgatgatgcca	gtgcttgccc	aaaaagaggg
49081	catcccttcg	aaataaaaag	gcaaagcttc	ttttgctttt	tatttttact	tctttctgct
49141	tgatagctag	gtataatatc	tagttttggg	gcagccatga	ggataacagt	tgacatttta
49201	ggatgggtga	gaaagttaga	aagaggctaa	atccttgata	acactgcaga	gctattgctt
49261	ctatcattag	cttaattcta	ggcttctcac	tgagtgaaga	aaatagacta	attgttgaaa
49321	ctattgcttt	catttgcttg	aagtcaaaata	tattccttct	ctgtgtgtcc	ttttgacaa
49381	atctttaaag	ttataaatac	atagaaatat	cacttttcac	ctttattctt	ttacaagtat
49441	taccttggtt	ccagaacatt	gatttgagt	aacaaatggt	taatgtttga	aaactttcct
49501	gaaatttctt	gcctctaaca	agagccagaa	ggaagactgt	tcacatattc	tttatatctt
49561	gagccagatt	gtcaattcct	ccaggccata	tccattactt	tgtgattatt	gattattcag
49621	tttatgaaaa	tgagtatgac	acaagttcct	tcttgcaaat	cagatgtcaa	tacctcagaa
49681	ggctccttcac	ccaaattatg	gcagccaccc	agtcattcac	aaccacggca	tcttgtttta
49741	atttttggt	tagccctgat	aaacctctga	tacttttcta	atttacttat	gtatgtattc
49801	acgcatgtgt	gcctttgttt	acttttagta	tcctccccc	agtgttagga	tatgagcaga
49861	gaatttgct	gctttatcta	gaaatataatc	cttggttat	aggagacaag	taggcatatt
49921	gtttgaatga	ctgaatgact	gaatcagtaa	gtatgagaat	caaagaattg	ttagatatag
49981	tcactgacct	gaggaaattt	caaagacaaa	taaggaaatg	agagacttac	tagaaaaat
50041	gaaaaaatga	agttataaga	atgctttaaa	aaaaaggcca	agtggtgtgg	ctcacatctg
50101	taatcccagc	actttgggag	gccaaggcag	gaggatcatt	tgaggtcagg	agtttgagac

50161	cagcctggcc	aacctggcga	aaccccatat	ctactaaaaa	tacaaaaatt	agccagacgt
50221	gggtgtgggt	gcctgtaatc	ccagctactc	aggagtctga	gacaggagaa	ctgcttgaac
50281	ccaggaggtg	gaggttgcag	tgagccaaga	tcgtaccact	gcactccagc	ctgggcaaca
50341	agcaagactc	tgtctctaaa	aaccataaaa	aggatattaa	aaggacagag	gtacagtcag
50401	aacccaatga	agataagcca	caacttggat	tttttagtag	ctcagaggac	caagactttc
50461	atatggattc	aggcagcaga	gagagcaaa	ccatagtga	aggaggagg	ctggagcaaa
50521	gcatgttgta	gcataacata	tctgtgagtt	taaccagtaa	aaaaggcttg	gtgtctcttt
50581	gagaacattt	gtgatgtggt	tggaaaaatg	ggagactctg	ctcaaagatt	gtgggttttag
50641	gattggctat	gctgagttat	taggtgagtg	caaaaagta	tgcagttttt	gcattgttgg
50701	aatttgccat	ttgatattgg	aataactctt	taaaataatg	tgcttatgtt	atacatcatt
50761	taatgggcac	ttctcacttt	ctattttttt	actaatgaca	ttacttgctg	tttattttgt
50821	gtttatttta	gacaacggaa	acaatgttag	ataaaaagca	aattctagcg	actttcttat
50881	ttgagttcaa	aatgggtcgt	aaagtaagca	gcaaagacaa	ttcacatcat	taacaaagca
50941	tttggcccgag	gaactgctaa	cgaacgtaca	gtgcagtggg	ggttcaagaa	gttttgcaaa
51001	ggagacgaga	gccttgaaga	tgaggagcgt	agtggccggc	catcagaagt	tgacagtgc
51061	caattgagag	caatcatcaa	agctgatgct	cttacaactg	catgaaaagt	tgctgaagg
51121	gtcaatgtcg	accattctac	gggtgttcag	catttgaagc	aaatcagaaa	gggtgaaaag
51181	ctcgataagt	gggtgtctga	taagctgacc	aaaaatgaaa	aaaattgtcg	ttttgaagt
51241	tctctctctc	ttattctaca	caacaataac	gaaccatttc	tcagttggat	tctgatgtgc
51301	aaggaaaagt	agatgttata	tgataaatgg	caatgaccag	ctcagtgggt	gggcccagaa
51361	gacattccaa	ggcacttccc	aaagcacaac	ttgcaccaat	aaaggtcatg	gtcctagttt
51421	gatgggtctg	tgccgggtctg	attcactaca	gctttctgaa	tcccagtaaa	accattactt
51481	ctgagaagta	cgctcagcgc	atcaatgaga	tgacacaaaa	actgcaacac	ctgcagttgg
51541	cattagtcaa	cagaaaagggc	ccaattctct	acaacagcca	actgcacgtc	acaccaccaa
51601	tacttcaaaa	gttgaacaaa	ttgaggtggg	aagtttttgc	tttatccacc	atattcacct
51661	gaccccttgc	caaccaacta	ccacttcttc	aagcatctca	acaacttttt	gcagggcaaa
51721	tgcttccaca	accagcagga	tgcaaaaaat	gctttccaag	gggtcgtcgg	atcccaagc
51781	atggattttt	atgctacagg	aataaacaaa	cttatttctc	cttggaacaa	ctgtgttgat
51841	tgtagtggtt	cctattttga	ttataaaaga	tgtatttgag	cctacttata	atgaattaaa
51901	attcacagtt	caaaaactga	attactttta	caccaaccta	atagtttata	cctagctttg
51961	ctgagcctca	agttccttat	ttgataaatg	gggatgagaa	tgcttacctc	gcatatttgt
52021	aatcaggatt	aaaagagaca	gtatgaaagc	aatgtgagaa	gcatgagtaa	ttatgaaaat
52081	ggctaacagc	tcttgagttc	ttattatgtg	tcaggcactg	gagtcatcac	ttcacacaga
52141	tgatctcttt	taatcctcag	aacaacccag	aggaaaagac	aatatccagg	tttcaacgtc
52201	aaagggagag	agccagctag	gagcagaact	gggctgggga	gggggtttgg	gtttgggtga
52261	tggaagaca	gctcttggga	aaacattcct	ttctttgacc	acctcatcta	acaattgctc
52321	tgatgatcca	cagttccacg	acttatttca	cacttaaact	aggaggatga	tatccaaacg
52381	tgccaccag	gtgagacata	tatatacatc	ccattttcac	caagcatcca	aggatctact
52441	ttggcaaaat	acattcaaca	acttggggta	attctgtaac	ttccttggag	aggaaagtga
52501	gcctcttcaa	atttactttt	tgcatccaat	tgttcacaag	ccaagtgtat	ttttaaagat
52561	acattttctt	aactgccatc	acatgcaaat	ttctaaggaa	attagaatgc	tgccaaacat
52621	cttccagaaa	tggaggtagt	aactcttgca	cagaaagaa	atgtaccata	agttagaaaa
52681	tctgaatag	aattccaaat	attccagtca	tgatttacc	aagcagaatg	gcataagaat
52741	taggagcata	gtctccaggc	tgacaaaatc	gtttactgcc	tggtgtcac	atattgttct
52801	gtaaacttgg	ggatgtttat	ttattcacca	tgtaaaatca	tttttctcaa	gtgtaaaatg
52861	gaaacaaaaa	aacgtattac	tcagtgcta	ctgtgacaat	taaatgagat	gatataatg
52921	aagcaaaagg	ttgtgtggtt	ctcagggaca	gtactgaaaa	tggtaaattt	ttttcttttt
52981	actactaatt	tggttaagtct	ttcagatttc	tcacttggtg	ctttagctat	ctttattatg
53041	tgctactcat	gacaaaattt	aggttttgat	gtaatttaag	tattggaaaa	tacaataagc
53101	atattacctc	tttgataatt	atccttttag	cacaagttag	cattcagtc	agccaaaatt
53161	acataatgct	aaatgctgaa	attagtttac	ataaaaagta	gtgggcccgg	cgcagtggct
53221	cacccctgta	atccaagcac	tttgggaggc	cgaggtgggc	ggatcaccag	atcaggagat
53281	cgagaccatc	ctggctaaca	cgtgaaact	ccaactctac	taaaaatata	aaaaatttag
53341	caggcgaggt	ggcgggcgcc	tgtagcccca	gctgctaggg	aggctggggc	aggagaatgg
53401	cgtgaacccg	ggaggcggag	cttgacagtga	gctgagatcg	tgccactgca	ctagagcctg
53461	ggcgacagag	ctagactcca	tctcaaaaaa	aaaaaaaatt	agtgttctgc	agactctgtg
53521	atatttgttg	tttagtatca	gttctcaaaa	tggagatagt	acatgtgata	tcaccagaa
53581	aactccagcc	aaacctacta	tttaatacta	aatagtaaaa	tagtcaatat	gtacactcag
53641	tacatttaca	tagttctaat	cagcatggga	tcattaagt	ttgatgttct	gttattcttt
53701	atggagattt	tggttttagt	ctcaagaagg	attttcaaaa	tcacttaagg	gattcattta
53761	ctcatgtttc	tgaatcaata	tatgcaaaaa	gattttctct	ctttaccaga	aaatttacct
53821	aatatattat	tctcttaatt	tgtgattctc	tatgcctgtg	ggaaaatata	aaagaaatct
53881	taataataatt	gttattctat	ggaaatgata	ttcatgtaca	tttcttgtgc	ttccctatgc
53941	gtaattttta	gagtggaatt	ggccccttat	ttaaactctg	agatttattt	gctctctctt

54001	aaagagacag	aagactccta	attggaaata	aaaattgtag	cctagatact	tattcaataa
54061	tttaatatga	cctgcgttgt	taagttcgaa	atttcactcc	agcattacca	gtgaatttta
54121	tattttatct	tatttagaat	ctatgttgtt	atatttttgg	tagttgtaat	catgcctgtt
54181	tccccagcta	gaatgcaaac	tcttttgagt	ttggaccatt	taattaagct	ttggaacatt
54241	cacccccaac	tcacaggtca	tctgtaacat	aacagatctc	gccgtccact	taaagtgggt
54301	gaatgaactt	cgtagggggt	aaattatgcc	tcaagaaggt	tgggattaca	ggcctcagcc
54361	accacaccag	gcctttgttg	tcgtatttga	ttgttccttc	catgtgcttg	aggctccact
54421	ggctctctgg	gtcattgtct	gcgtccttac	tcattttatg	ccagagaagg	taaagagaga
54481	tatctggtaa	atttggaata	aaattatctg	ggcctgtgta	ttagttcatt	ctcacactgc
54541	tgtaaagaaa	tacttgagac	tgggtaattg	ataaagaaaa	gaggtttaat	ggctcacggg
54601	tctgcaggct	gtacaggaag	catagagcgt	tctgcttcgg	gggaggcctc	aggagctttt
54661	tactcatggc	ggaaggcaaa	gcgagagcat	ggctggagca	ggaggaggag	agagagagtg
54721	gggaggtgct	acaccctttg	tttttgagac	ggaatttcga	agtttcactc	ttttcccagg
54781	ctggagggca	atggcgctat	ctcggctcac	tacaacctct	gcctcccagc	ttcaagtgat
54841	tctcctacct	cagcctcccc	agtagctggg	attacaggcc	cccaccacca	cgctcggcta
54901	atttttgtat	ttttagtaga	gatggggttt	caccatgttg	gctaggctgg	tctcaaactg
54961	ctgacctcag	gagatccatc	cacctcagcc	tcccaaagtg	ctgggattac	tggcgtgagt
55021	acacactttt	aaacaagcag	atctcttgag	aactccatca	tgagaacagc	actaggggga
55081	tgggtgctaaa	ccattcctga	gggaccaccc	ccaggaccca	atcacctccc	accaggcccc
55141	attttccaata	ttggggatta	caattcaaca	tgagatttgg	acaggcacag	agatccaaac
55201	catgtcagcc	tagttctact	aatcctggcc	ttcttttatt	ggaattctcc	aggtccttgc
55261	tgtctctctg	agcgaggaaa	gacctgttct	aYaataatcc	agcactgcag	gctggactaa
55321	tcactctttc	aacacgtggg	tcacttagca	agcctatttg	gcctcacaca	tcctacactt
55381	agctcttact	agtaactgaa	atggcacttt	gccttctgct	tgccactcct	tatttttcac
55441	aaatgttcac	aaatccaggg	gaagaaaaga	ggtgctcttt	ggtgagcttc	cttggaataa
55501	ctgattctat	acttttattt	ctcactttgg	taattcttat	ggtgtggagg	ttcattgtacc
55561	ttcagagttt	atatgctgaa	atcctaacc	ccagggtgat	gatgtaagga	ggaggggtct
55621	ttggaagggt	atgagatcat	gggagtaaat	ttctgtcttt	aaaagctatc	agtttatggc
55681	attttgttat	agcttcccaa	ataatgaaac	ccaatagcta	ctgtgaatta	agtagtatac
55741	ttaacctcca	ttttattcca	ataaatgtaa	tcatttgact	cccttctatt	aagccctcac
55801	ctctacgtgc	acattttatc	tctcccactt	acctctcacc	ttccatcagt	ttccttctta
55861	attttacatt	ataatgccta	ttaacgttta	tgttctgttc	tataatcata	attaaatatt
55921	tatgctatct	ctgtggaggg	attctaagag	ctgaaaacta	gtaaacagca	gttacagtat
55981	aatgactaca	taaataattca	atgaagggcc	aactcacgtg	ctactattct	atttccttca
56041	ctttattccc	atatcaccca	ctcccatagc	aatcaacgtg	taatgatcct	aagcatcctg
56101	gtcaaatgat	tctccataat	tatacttcaa	tatttgctca	aaatcatgcc	ttatttaatg
56161	tatctcacac	ttggacacta	actttataga	cagggttagc	tgttattttt	cctagagttc
56221	atatttgtct	tccaggcatt	tgtatttatc	cttactgaga	gaagaagcat	attatttctt
56281	cactgtgtca	ctactgtaaa	cctgctcatt	tgtcattcga	tttattaatt	ggaatgccca
56341	ttcatctttc	tgagaacctc	cttgcagcct	actaacctca	agttccaatt	tcccatttgg
56401	tactttctgg	gccactgtgc	cactgtttat	gctgagattt	cttttgactg	acttctagg
56461	tcagctcagc	tggtttttYt	acctgttatt	ctctgccact	caaacactgt	catttgatca
56521	gtctctatgc	atttaccag	tatttactat	gtacaaggca	tagtattata	tactatgaaa
56581	attcaaggca	aatcaaagtt	tgttctgcac	caaaagagtt	aatagggttaa	gtcttgaaca
56641	cacatataaa	taatcataac	ataaaaacta	caaggaagat	aacagagaaa	tagaatttga
56701	caaagaaatt	agaatgaaag	aaatagaaat	atgccataga	tgaaaaMttt	aggagagcat
56761	cagatgtgaa	ggggtgaagg	ggagataggg	catggtgaga	gagagtcaag	gaaggctcaga
56821	gaatgcatga	gaaagaaaaga	agcaatctca	ccttggttaag	caacggttgt	gcaaaaaacat
56881	agacactttc	atagtgaaca	aaatgttaca	ctgctattct	acatagtatt	ggaagatctg
56941	gccagggcaa	tcaggcaaga	gaaagaaata	aagggtattc	gaataggaag	ataggaagtc
57001	aaattatctc	tgtttgcaga	ttacatgatt	gtatatttag	aaaaccccat	tgtcttagcc
57061	caaaaacttg	ttaaagctgag	aagaaacttc	agctaagtct	caggatacaa	aatcaatgtg
57121	caaaaatcac	aagcattcct	atagccta	aatagacaaa	cagccaaatc	atgagcaaac
57181	tcccattcac	aattgctaca	aagagaataa	aatacctagg	aatacaactt	gcaagggatg
57241	tgaaggacct	cttcaaggac	aactacaaac	cattgcttag	ggaagtaaga	gaggacacaa
57301	acaaatggaa	gaacattcca	tgctcatgga	taggaagaat	caatatcggt	aaaatggcca
57361	tgtgtcccaa	agtaatttat	agattaaatg	ctatttccat	caagctacca	ttgactttct
57421	tcagagaatt	agaaaaaaaa	aaaactactt	tgaatttcat	atggaacaa	aacaaaaaaa
57481	agagctcgta	cagccaagac	aatcctaagc	aaaaagaaca	aagctggagg	catcacacta
57541	cctgactttc	aaccaaaaga	gcattggtact	ggtacaaaaa	cagatatata	gaccatgga
57601	ctgacacaga	ggtctcagaa	ataacaccac	acatctacaa	ccatctgac	tttgacaaac
57661	ctgacaaaaa	caaacaatgg	ggaacaat	ccctatttaa	taaatggtgt	tgggaaaact
57721	ggctagccgt	attcagaaaa	ctgaaactgg	accccttcct	tacaccttat	acgaaaatta
57781	attcaagatg	gattaaagat	ttaaacataa	gacctaaaac	cataaaaaac	ctagaagaaa

57841	acctgggcaa	taccattcag	gacataggca	tgggcaagga	cttcatgact	aaaacaccaa
57901	aagcaattgc	aacaaaagca	aaaattgaca	aatacaatct	aattaaagag	cttctggaca
57961	gcaaaagaaa	ctatcatcag	agtgaacagg	caagctacag	aatgggagaa	aatttttgca
58021	atgcatcttt	ctgacaaagg	gctaataatct	agaacctaca	aggaacttaa	acaagtttag
58081	aagaaaaaaa	caaacaaccc	catcaaaaag	tgggtgaagg	atatgagcag	acacttttca
58141	aaagaagaca	tttatgtgac	caacaaacac	ttgaaaaaaa	gctcatcatc	actggtcatt
58201	aaagaaatgc	aaatcaaaac	cacagtgaga	tatcatctca	tgccagttag	aatggcaatc
58261	attaaaaagt	caggaaacaa	cagatgctgg	agaggatgtg	gagaaatagg	aacactttta
58321	cactgttggt	gggagtgtaa	attagttcaa	ccattgtgga	acacagtgtg	gcgattcctc
58381	aaggatctag	agctagaaat	accatttgac	ccagcaatcc	cattactggg	tatatacca
58441	aaggattata	aatcattcta	cataaaagac	acatgcacac	gtatgtttat	tgcggaactg
58501	ttcacaatag	caaagacttg	gaaccaaccc	agatgcccat	cagtgttaga	ctggataaag
58561	aaaatgtggc	aggtatacat	cgtggaatgc	tatgcagcca	taaaaaagaa	tgagttcgtg
58621	tcctttgcag	gaacacgagt	gaagctggaa	accatcatct	tcagcaaaact	aacacaggaa
58681	cagaaaaata	aacaccgcat	gttctcactc	acaagtgaga	gttgaacaat	gagaacatat
58741	gggcacaggg	aggggaacat	cacacactgg	agcccgtcag	gggggtgggg	ccaaggggag
58801	ggatagcatt	aggagaaata	cctaattgtag	atgacagggt	gatgggtgta	gcaaaccatt
58861	atggcacatg	tatacctatg	taacaaacct	gcacattctg	cacatgtatc	ctagaactta
58921	aagtattata	ataataataa	aaaaatccct	ttgccttctg	ccatgattgt	gaggcctccc
58981	cagacatgtg	gaattgtaag	tccattaaac	ttctttttct	ttataaatta	aaaaaaaaag
59041	tacagtactc	ttagtctaaa	gagaaactaa	aaccaatttc	atccattttag	aaaaaataac
59101	aactgaatat	tgaaaacccat	gggacactgc	caagatcatt	ctcagaagaa	aatccatagc
59161	aataaatgtt	gtattattac	ggagcaataa	aacctgggag	aaaaaaaaatga	actaagcact
59221	aaccacaagc	acactgaaaa	atatctaaaa	aataaatctg	atctgtgaga	catataaatt
59281	gatgaatata	aactttaaat	ataatgagtt	agacatctaa	ttaaaaagtt	gaattcaaac
59341	ataaaaacca	gagtggtttt	ttaggaaagt	taattggaata	gatagatttt	gtggcacttt
59401	ttttttaaga	acaagaaaag	tcatttatatg	gtgaaaagata	cagttactat	aacacatata
59461	atattataag	aaagtagtgt	atataaatat	cacatagaaa	acttgaaacc	tcgataaaat
59521	gattgaattg	tttgatatata	tataattatta	aagttgctta	tatatcagaa	taaactaata
59581	actacaaaaa	aaatgttttt	agtgtaaggg	aacaccctct	gcccctccaa	aaagggcaag
59641	aaattcaagt	atttctatat	aaagaaatatt	gctttaaatg	tcttaggatt	caggtaatga
59701	tgatgcagac	agaagagggtg	cagactaatc	taacttccta	ctaaaaccca	aaatctaaaa
59761	taaaaaatat	agataaatca	aatctaattg	agcaggaggc	ctgttaaat	ttattctaag
59821	aatgaaattt	atttatattga	cagggataat	aaaaggagaa	aggagataa	aggaatacaa
59881	gatgattatc	tggatgttaa	tttgggatac	ataatgaaga	gtaacattat	tcattgagat
59941	aagaaagaca	gcttcgtgta	gattgaaata	tgttgcttga	atttagatat	aaagtttggg
60001	atataatata	aaatttggag	atataattac	agtataaatc	cgagagtcac	aagcatttagc
60061	aaaatcaact	agagagagcc	ttcagaattc	taacagaaca	gaataagata	attaacttta
60121	caaacatat	tttgaatatt	gcattaaacca	tctattgcta	tatatataga	caaattacct
60181	caaaaattcat	tcagcagttt	aaaactcaag	aataaactat	aattattcct	tgcatagttt
60241	ctgagggtca	ggcacctggg	agaactcttag	ctgtgctatt	ttggcccaga	ttctctcatg
60301	atattgcagc	caggttggtg	gccaggactg	ctttagccat	cgcaaatatt	ttccggggcc
60361	agggtatcta	cttccaagct	cactcgtggc	ttttcttgca	ggcttcagtt	cctcaccaaa
60421	tggacctcgt	catagtgcg	ctcacaaaat	ttaagtttgt	tatccccagc	ataaaggaa
60481	taagagaagg	tgagagagat	atacacaaag	tggaagtaa	atcttataac	ctaattttgg
60541	aagtataacc	caatcatttc	ttctatatatt	tatttcttag	aggcaagtta	ctaaattcag
60601	ccactcaacc	agtaggaaaK	ttaagctctt	actgctcaaa	aagagtatca	acaaatatgc
60661	caacgtatct	ttaaaaccct	cacaaatMcc	aaccacttgg	aagagctaag	aatttgtgat
60721	gagaaccaag	acaaaactat	taggaatcta	cataaacatg	aaggcgtag	tgatgtcctg
60781	gaagtgcagt	gatttgaatg	agccactgtg	tgtaagata	gtggaagaa	ctattttttt
60841	ctgagtgcc	actatatacc	tggcactgta	ctaggtgtgt	tatgcgaatc	atctcttttt
60901	cttgccatta	ataacaacca	tgtacttcca	gtgccttcac	cacctggaat	cccaagggtc
60961	ctgaatcaga	taggataaat	tttggttttg	aaggacagag	tctcaagata	gtaggaatta
61021	aaacacaatg	gaacttatct	ttctcacaca	taaaaatMtg	gaagtaggca	gactaacgta
61081	atgggagcag	tggctctgct	ccatgaaatc	ttggggcatt	ccagcccaYg	tttccatcat
61141	tgatagccca	cagtgctatg	atgcagaaca	cgtctagagc	tccatctttc	aagaagcaca
61201	atggagggaa	gattaagacg	aaaggagcaa	aggacagggtg	ccagctgtct	tttaggaggc
61261	tcaacaggaa	attaccaaac	acttctactt	acctgtcaca	acttttccat	ggactgagag
61321	atctctttac	tctggtagta	ataaagccca	gataaaatct	aggggcttta	ttactatgga
61381	agaatagata	tatgggggta	aatagtagtc	tttatgacag	tatcatatgg	aattaccaga
61441	agcattctaa	agtacaggac	ctctgaaaaa	attattagcc	ttttgattcc	tccaggcatt
61501	gtgaaatctg	ttagttttctc	aagctcaagt	ccttggacta	attttgacca	agttttatat
61561	agttttatatt	cttagtYgaa	gtacacaatt	ctatttcatt	tctaccctaa	cagctcacat
61621	tatcatgaat	gagaagaaat	gtgcacctgt	attttaaaat	tttttgccca	atthagcagt

61681	cttgggttctc	agatgtccca	tctagtaata	atatcttctt	attggctttt	gctattttgt
61741	gggctagaga	aaccacttgg	aaggaccata	aaagttccca	gttttatgtt	ggtactaaat
61801	atattctttg	catccattaa	tgtctcacc	aaatttttta	tcatcaatac	aatgtacaaa
61861	attatgtatg	gtgtgattat	ctgtgatgtt	ctcaaatccc	acaaataatt	attagtaaac
61921	aaggactaga	gcccagggtg	gagaattgta	attattttatt	Rctaataattc	ttagacaaaa
61981	gggaatgcca	tttttcaacta	ctcacacata	aaaaaatgca	gtatataagt	aaatgtttta
62041	gcgctgtcca	aaaaccacta	tgagttttgg	ggctcagaaa	attatacccc	aaagtgaagg
62101	cctcagaagt	agactcagaa	gcaaagtttc	tctgaacttc	tgcccgctg	tctctcgac
62161	ctcattcttc	cccccaatgc	aagtcataga	aacttgaatt	cttcctttcc	aagggtgggtc
62221	atagaaaccg	aacccctttt	ccccaaagcc	agccataaaa	cctgaaaaca	ttaccctaac
62281	cttccccctc	ctttctgtgt	aaaagatggt	cataaaagaa	ttctctgacc	tgacttgttt
62341	gactgtaggt	cgtaagaccc	cctgagttcc	gaaaaacactt	cataccccg	aggaaggaat
62401	gctgtactga	gaggccaaga	agcatctgaa	cagacagacc	atgctgagtt	tccccaat
62461	gtctattccc	atcagggtcat	acccttttgt	ccaaccatat	ttctacacg	tttttcattc
62521	ttcatcaa	ctaagcataa	aaatagacag	ttgtcactgt	atctttgagt	ctccattttt
62581	aaggctccca	tgacacataa	aactataatt	aaataaattt	gttatgcttt	tctctgttta
62641	atttgtcttt	tggtatagga	gtgtcagctg	tgacccttat	gatagggagg	aaagggatca
62701	cctcctttgt	gcccctacaa	gagaaaaaca	caggtcaaag	taggctgaga	gcatatttct
62761	gactcttaag	ctcagactta	tggttaccac	caattctatt	tgttcccat	tttaaacc
62821	tctcaaaaat	attttatcct	cactgagcct	cagaaagatt	aaatatcatg	tccacactta
62881	catactaat	aaataccagg	attgggattg	taaattcatg	tcttcttaac	gtcaaaacac
62941	atcttcatat	tgtttccaac	acaaggagac	ttgaatgcaa	ttactttttt	tttctgtttt
63001	ctttggtata	tataatttaa	gagcattgtc	attgatctta	gggtgatttt	ttcttataaa
63061	gtatttggtta	aatttcattt	tattgacatt	tgaataaatt	gatgccaaag	tattgcattc
63121	aacatttgct	ttaaaaagta	tattgtttgt	tggtgtttgg	tcatttttaa	atgtaattgc
63181	aacataattaa	aatggggatt	atctagataa	ttaaactgga	tctctgtaat	gtgacatacg
63241	tgtgtcccca	catatagcta	tcctcaatc	caaagaagag	agggaaat	ggacactgaa
63301	ttattctgcc	agctgactcc	gcaccccttg	gtgactgata	atcttctata	Mgttgactt
63361	ggcagtgcta	gcaagtaagg	attaatatcat	tttaatcctc	cctgatttgg	tccttttttt
63421	tttttgtagc	tttaactattt	tccagtcctc	aagactttct	attgtctatg	gttaatttaa
63481	acaaatcata	ttttcacaga	tcctatggat	ctttctagag	tttttttaag	tttctactg
63541	tagtttattt	tcaaatatat	tttaacatgg	tgacagttct	ttcacttgct	tttccctttc
63601	attgcatttg	cttccaggct	aatctcctga	agtacaattt	gatcatgttt	tccgatataa
63661	cgttgaaaga	gctgtaagtc	atatcagacc	ttttagtaat	ctatacagtt	ccagaatcat
63721	taaggcaagt	gagttagaaa	aaaacatttt	caatgggata	tttatgtttc	atgactttag
63781	agagaatttt	aaaacaagaa	aatttcactt	tttatcaaac	cctctcaata	ttgtttttta
63841	tctcatagat	aacttttcac	ttattttaat	tcaattttta	aattttatta	ttttcatagg
63901	aatttttagga	gtactgtaga	acacaYtgaa	gttcctgact	tctttttaca	gggtgaatta
63961	ctgagattta	ggttctatgg	ttggcccaag	tttcataaat	aataaatgga	aaataaggac
64021	ctaaagccta	tagagtccaa	tttctatgta	cagaatctct	ctgtaatctt	tgctctcttt
64081	gtacacatta	agggatcagg	agatgctaag	attgctaaat	acacagtagt	tttctaatat
64141	aaataatgat	tagatataca	tagaaatatt	tttaagcaat	tcagatcctg	cccaatctct
64201	ctactccctg	ttgctaccat	ctaagatata	tctactgcct	ttgaatctgt	cagagtgaga
64261	ctcagcaatt	cataactagg	gcactctgag	acagctggaa	cttccttaga	gaaccagctc
64321	gatccactga	accattggca	agatgaattt	tcctacagca	tcatacgta	gtgttattcg
64381	tgttcccaac	cccacatccc	taggccaggg	gtcatatttc	taaagtcccc	tgtttatcta
64441	ataatgtatt	ataaacttga	aatttgctaa	gtatagatct	taagtgttct	caccacacat
64501	acacacacaa	aagtgtaaat	atgtgtgata	tgctgtatta	gcctgttttt	acgtgtctga
64561	taaagacata	cctgagatgg	caatttgcaa	aagaaagagg	tttaatggac	tcacagttcc
64621	acgtggctaa	ggaggcctca	caataatggc	agaaggtgaa	aggcacttct	cacatgggtg
64681	cagacaagag	aagagagctt	gtgcaaggga	acccctcttt	ataaaacat	cagatctcat
64741	gaggcttatt	cactattaca	agaacagcac	aggaaagacc	tgcccccattg	attcaattac
64801	ctcccactgg	gtccctccca	tgacatgtgg	gaattgtgga	agccacaatt	caagggtgaca
64861	tttgggtggg	gacagagcca	agccatatca	ttccaccctg	ccccctccca	aatttcatgt
64921	ccttaccatt	caaaattaat	catgccttcc	cacactcccc	taaagtctta	actcatttca
64981	acactaactc	aaaaaccag	aatccaaagt	ctcatttgag	acaaggcaac	tccttccac
65041	ctatgagcct	gtaaaatcaa	aagaaagcta	attacttcct	agatatataa	tatctatata
65101	tctatataga	tacataatat	ctatatatct	gtatagatac	ataatatcta	tatatctgta
65161	tagatacata	atatctatat	atctgtatag	atacataata	tctatatatc	tgtatagata
65221	cataatatct	atatatctgt	atagatacat	aatatctata	tatctgtata	gatacataat
65281	atctatatat	ctgtatagat	acataaatatc	tatatatctg	tatagataca	taatatctat
65341	atatctgtat	agatacataa	tatctatata	tctgtataga	tacataatat	ctatatatct
65401	gtatagatac	ataatatcta	tatatctgta	tagatacata	atatctatat	atctgtatag
65461	atacataata	tctatatatc	tgtatagata	cataatatct	atatatctgt	atagatacat

65521	aatatctata	tatctgtata	gatacataat	atctatatat	ctgtatagat	acataatata
65581	taattttataa	attatatatt	atattttaatt	aaattataat	ttaattataa	tttaattaaa
65641	tataatataat	aattttataaa	ttatatatta	taaattatca	cctgataatt	tataaagcca
65701	tcagatctca	tgagactttt	tcactattac	aacaacagca	tgggaaagac	ctgccccac
65761	gattcaatta	cctccactg	ggtccctccc	ataacacRtc	agaattgtgg	gagttacaat
65821	tcaaggtgac	atgttggtgg	ggatacagcc	aaaccatata	acatgttaat	tagcttgatt
65881	cttataatta	tttcaactata	catatacata	gtgaatatat	atataatgt	atacacacac
65941	acacacaaac	acacacaatt	ttatctgaat	catacctcaa	taaagctgga	aaatagatcc
66001	caaccaaggg	ccaggtatag	aagcatgtgc	ttataacaat	aaMRtttcta	aaagaaatag
66061	ccccgcttag	gtaaagtgtta	ccaattttctc	taatacccca	aaaattcaga	tttttgtgtt
66121	aaaaatcctg	atttttaata	ttggtaaaac	ttttaataac	atagataagc	aatttatgtt
66181	ctttatggat	ataggggtccc	atgggctata	gcttacaata	tgtgggtctca	atctcagata
66241	ttctaYagtc	ttcagagtct	ctcatgttta	gacctgaatc	atgtctccag	tttctctaaa
66301	attgagtttag	tttaagaYgt	gatcccaaag	ctcaccctag	gttgccaggt	accaggttcc
66361	aataggtgaa	tgtagagttc	tacggatgag	ggagctgttt	gatttatgca	tactcatgtc
66421	agtttttga	ttatgccaac	tctgaattca	gtagtttggg	atacattcgc	aagaaaaaat
66481	gaaactcata	tgaaagattt	gtcaattata	aagcctcaaa	agtggggcat	aattaatcca
66541	tgtttagKgt	tttggatttg	tttcattctg	gaataaggta	taacacatgt	tgtcttgagg
66601	atttttcttt	tcctttcaaa	cataaaagta	aagttgggtat	tggttttagta	atagcacaga
66661	cacttaaaaa	gttaggcttt	tattattctt	tttgtgagtt	taaaatagga	aattctcagt
66721	ccaaaaaat	gggcacatc	tctctgcca	agtacaaact	gtttggtaga	ttgatatatg
66781	atagctcaat	acagagaagc	aaaccagtct	cttccctaaa	tctgtgtgat	cagaggttga
66841	acctaaatca	ccacctctgt	tatttgccag	atatcagcac	aatggcacia	tggacatttc
66901	acacatacca	tgaacccaag	ctcattgcaa	ttatttacta	tcaaattgacc	attacatatg
66961	tttcctcaaga	tccgttccca	aaatatgttt	ttcataaact	agtctccaat	tatatacatg
67021	ttactttcta	gttatttttt	ccattttttta	tttttgattt	ttgtgtgtac	atagtaggtt
67081	tatacatttt	tgaggcacat	gtgatgttct	gatacaggca	tgcaatgtga	aataatcgca
67141	ttatggagac	tgtggtattc	atccctcat	gcatttacc	tttgtgtac	aaacaattaa
67201	attacactct	ttttgttatt	tttaaatgta	caattaagtt	actattgatt	atagtcagt
67261	tggttggtcta	tcaaacagta	tgtcttattc	gttctttcta	attttttgta	ccactaacia
67321	actccacttc	cctctcaatc	tccactact	ttttccagcc	tctggtaatg	atccttctgc
67381	tccctgtgtc	tatgagttca	attgttttga	tttttagatc	ccacaaataa	gtgagaacat
67441	gcaacgtttg	tctttctgtg	cctggattat	ttcatttaac	ataatgatct	tcagttccat
67501	ccatgtgtct	agttacttta	ataaacacaa	acaccaacac	aaagaagcaa	atggaaaaca
67561	tcagtacaaa	taatcttccc	tggccttcat	aatacactca	aacactgatg	aaagccagat
67621	agcttttttc	catataattc	aaaataagaa	gtaaaagaca	tttttaaatg	aaactaaata
67681	tatgcatttt	tttttcaaaa	ttttacttgt	cagtctcttc	tcaataacta	gaatgtaaat
67741	ttcctaaatg	tttttcttta	tgatagcatc	ctctaaaata	ttttgcctat	cgcagaagtt
67801	caacaaagat	ttgtgaattg	atcagtcatt	aggctctaaca	ggataaataa	ctctgactta
67861	aagccaattt	aagccctaag	gcaatatggg	ttttggttca	ttacaaaagc	tttttcaaaa
67921	aatgcataga	aactctttct	aaaaaatttt	ctaatttgac	atatccagag	gggcaccatc
67981	cttttgccat	ggattttatg	ccaatgtgtc	ttcctgaagt	ggacaatgat	ttatacttca
68041	catatctaca	agttttcaaa	tagctataga	ttctcaagac	aaggtctgaa	agaagaatat
68101	aatcagttta	aaaatgttca	aaatttagaa	gataaaataa	aagctccaga	ggataaagag
68161	atcaaaagta	agaagggtaa	attgcttata	gtcgaacaat	taaagaaagg	ggctgtttaa
68221	aagtttggtg	cacaggccgg	gtgcgatggc	tgactcctgt	aatctcagca	ccttgagagg
68281	ccatgggtgg	ccgatcacct	gaggtcaaga	gttcaagacc	aggctggcaa	aaccccatct
68341	ctactaaaaa	tacaaaaatt	agcagggtgt	ggtggcatgt	gtctgtaatc	ctagctactt
68401	gggaagctga	ggcaggagaa	tcgcttggac	ccaggaggca	aaggttacag	tgagccaagg
68461	tcatgccatt	gcactccagc	ctgggtgaca	aagcaagact	ctgtctcaaa	aaaaagaaaa
68521	aaaatgtttg	tttcacagag	gtcaccatgc	aagaagattg	ggaggaggtc	agggttgaaa
68581	ggaataacac	atcagctagg	acagcacgtt	aagtaagggt	tccacatatg	tggttaagctc
68641	actcaggatg	gtggcagcat	tgaacttaga	gatgataatt	gtgaaccatg	tgcaccttcc
68701	ataaatgaag	ggaaaatacc	aataaacttga	tagatggcat	ctaataaggg	cacagaagca
68761	tgaaagagat	ggtgagtacg	ctgcagactc	cagtctcttg	gctaaatgat	gtcttctaag
68821	cagtgtgtgt	aaggctggaa	gatctaaaga	gtgttcaag	aatagattga	taactgaatg
68881	gagtttaactt	cagagagtgc	aagaattaga	agggtggagag	agagaaacta	gagatggggg
68941	cagggattat	atgcagttaa	gaggaagaag	tgaggggggtc	gatgaagaac	aagatggaac
69001	acctaaagcc	tcagacaagt	gaagtgaatt	attttatacc	atgaaggtaa	ttcagaacag
69061	aatccataca	aaaatgcacg	ttttctgtct	tcttaggccc	aactgttttt	ggccaaatat
69121	atagcaaaat	tggtggtgga	aaagcgtatg	tctcagaaaa	tttaagcta	aattgtcttg
69181	ttatgggtggt	tgattttttaa	tgtgagatta	aagtaaaatc	caaaatatac	tttaagtatta
69241	agattatgta	actagaaata	tttcttctc	tcttataatc	ttaaactcta	aacaatttca
69301	gcactctgact	acagcctgaa	taaatgcaa	tggaaacata	acaaattaac	aaagagtaac

69361	tgggaaatgt	caacattccc	agatgctgat	ggtcagacta	atgacgccaa	ttaagaaaat
69421	tgatcttgaa	aataaaacag	atgtccta	acagagtggg	agcacaattc	atgaagattt
69481	ggagcgccag	gagtagagga	agggaaatgg	acagttaagt	aagtgtttgg	tgagtccatc
69541	cagcctcaca	taattcagtg	tagataaagt	gcgatgtgct	aagaattaga	atggaaatgt
69601	acagaaaatg	ttaacaggac	agaggagaag	gagctcccaa	ctctgcaactg	aggagacaga
69661	agaagagttg	attcaacaga	tatttattgc	acacctatta	taaccaggta	ccatgtctggg
69721	cactagtga	cagttagcaa	tagattcaaa	acttctctgcc	tttgtgagct	gacattctat
69781	aagggagata	cagagaaaat	atataatatg	tcacatggta	acaagtaata	tgagaataaa
69841	aaaattagag	taaggaaata	agaatatgca	ttaaacaggg	tggttaggaa	atgtcctctg
69901	ctaaggtgac	acttaagcag	aatctgacat	ttgaggtgac	atatgacttg	gttctaaaag
69961	cgtgtattaa	gagactgcca	gaaagacaag	aggggaattt	ccagacagag	aaagtaaaaa
70021	atggaaaaac	ccagggactc	aaaaaaaaaa	attgggggaa	ataagttgtt	tttgaccccc
70081	taccatctat	tctctcttct	aatggtggta	ttagatgagt	agacaaacct	tacttctctc
70141	ctcttagagt	cccagctagc	tctgagaatg	aaattaatat	agcatagatt	aacaggagga
70201	aagcgtactc	tacaagtttt	acatggcatg	agagccctca	tagagaaatg	aagactcgaa
70261	ggagcagttg	gagtcagtta	cttacaatac	gaattggaca	aagagcagta	aacttgaaaa
70321	tgtagactaca	ttatatggg	aggcttaaaa	gataagagtt	attctaataga	aaattgtaca
70381	gaattctttc	agtctcaact	tctgttctct	gaagataagg	atgtcacctt	tctttctagt
70441	aaagagagtc	tttcatggga	atttcacctt	tttaagaaaa	agcccaagg	tcaaagtgat
70501	cttgcaacttg	ctgtctttca	agtgtcttca	acttaaatg	tcaatatgcg	ccaaaatagc
70561	atattttagc	cacttcggta	ccatgaaatt	cctttagaaa	actgtctcct	tcttctcagt
70621	ccatgtgttt	tagaagagta	aaacacgtac	ctaggcctga	ctaattaaag	catcaaattc
70681	ctgtacaatg	atgagtaact	caacattggt	caaacatggg	aaatccctgg	gcttgtatga
70741	aggttgctcag	aaatctctac	tgggtagccc	agtggtttcc	atgatataag	tctggagctt
70801	ctggaagctg	tcatcctaga	accagggtat	gaaaaataaga	caccacaaca	tagagctaaa
70861	atacaaaaag	agtaagtcct	gatcttcact	tttgtttttt	ttttatttat	ttattataga
70921	gacagggtct	cctctctctg	ttgtccaggc	tggagtgtctg	tagtgcattc	tcaactcact
70981	gcaaaactcca	cctccccatt	caaataattc	tcctgcctca	gcctctggac	tagctgggac
71041	tacaggcacg	ccactatgtc	cagctaattg	ttgtattttt	agtagagaca	gggttttacc
71101	atgttgccca	gcctgggtctc	gaactcttgg	cttcaagtga	tccagtcagt	caccttggct
71161	tcccaaagtg	ctgggattac	cagctgtgagc	cactgcaccc	ggcctgatct	tatttttcga
71221	aatcctagat	aaaaccaYgt	ttgcaggtaa	acttggccag	tatttttttaa	atctgtgggc
71281	caRtaaaactc	ctctcttttaa	cattagactg	gatatcctat	tatttggcaa	tcagagtctt
71341	gatagaatga	atatattgtg	ttggtcttag	tcagtttggg	tttctacaac	aaattattgt
71401	agatgggtggc	gggagggggg	gtgctcaaat	aattaacatt	tattgtttag	tcagatgctt
71461	ggtaggggcc	ctcttctctg	tcgtgtcttc	aYgtggcctt	tccttggcgc	atgcacagag
71521	agagaacgac	agagagagat	gctgtgtctc	ctcctcccat	tattgcaagg	gcacttgcaa
71581	tatgggagtc	cttgacataag	gactccttct	tcataactta	attgagtcct	aaaggcccta
71641	tctccaaata	tcatcacatt	ggggattaga	gtttcaacgt	attaatttcg	gggaacacag
71701	gcattctata	gcaacgttca	agaaaagag	tgtcaaatgt	tatttttagaa	attttgatgg
71761	cacttgggtg	agtatcccac	aaagacctct	aggaaaactt	tattcttggc	atttaaagaa
71821	agttagattg	tagaagcact	gagaaagtgg	gggaggtagt	ttcagaatat	agagtaaaat
71881	gttgttagga	gtttagaatg	cccacacaca	gaaccttgag	aaatcacctt	ctctggattg
71941	cttactcata	aaaatgtgca	catgtgttaa	gcacagtagg	tactgtggtg	agcagcaatc
72001	tgtgggtgca	aaaatagatt	taacatttat	acattgttaa	agagaaccag	agcctaacag
72061	tggttaaatc	agtaaaaaaca	gattagcggg	aaagagatct	cagtatagaa	ctgggctcaa
72121	ttttgaatac	gaaaaggaaa	aatagggtat	tatagccaag	gagcaggttg	gggtatgggt
72181	ggatggaaag	ttactaaaag	gtgacatcac	aggtaaaggga	aattcttgc	aaattgacct
72241	aacaggattc	ttgctaaatg	taggccagga	tgaYtgga	tcaaggtgga	gaaatgagga
72301	atttcatcag	atctgaaagg	tttgggattc	cagctaccac	tgggctatgg	aggcccagca
72361	aggatggagg	cccagcaagg	atggaggcca	aggttgagat	ctagtgaaga	agaggattca
72421	aggggtctga	ctaaagtgtg	ggcaaagaaa	gaatctttgt	cagtgtgacc	acaaatttta
72481	atgtgtttac	aatcacttct	tggtaaagcta	gtattgccat	gtttaaataa	catgagagct
72541	taggggacat	tgctcacggg	aaaaatgact	gtgcacttca	gaatcttgaa	actttgttgt
72601	tttgaactca	actaaaccca	cactttttct	tttctctatg	gatttactat	tatcatcacc
72661	aaatactaac	ctatgtaaaa	ttataaatat	ttataacatt	gcacataatg	ctttacatct
72721	acaagtctac	ttaaatgggt	ctcagggtat	gctgtactta	aagtatgggtc	aaaaagcacc
72781	tcatcagaat	cacataggag	gcttggttaa	agttagattt	ctggacctta	taaacaaact
72841	ggatcaaaat	ttctgatagt	ggtgtctccag	tataggcacc	actctagaat	attttaataa
72901	aaattcaagt	ctagtcttat	gcacactgaa	gtttgagaac	tattcaacta	tggagcagag
72961	taaatattgg	ttagatttag	acaaaacaat	gagactggga	tttattgata	cccactatgc
73021	tctttacatg	ctttatacac	ctagtctcat	tttaattatca	aaacaggaca	gtgagataat
73081	attttcccta	ttttacagat	gctgagatta	aatgagaaat	ggttacatca	tgcaactcaa
73141	tccacacatc	tgggcactgc	ccaagctgat	agggaaattca	ggtctatctg	atgcttatgt

73201	aatagtccaa	tgggttcacc	ttgcccgcgtg	cctagacaga	gctaatttat	caagtcaagg
73261	gaattgtaat	ggagaaagag	taattcacac	agagcctgct	gcgaggaga	cagagtttta
73321	ttattactca	aatcagtcctc	cataagcatt	cagggatcag	agttttttaa	gataatttgg
73381	caggtagagg	cttgggaagt	gggaagtgc	gattggtcag	gttgagatg	gaatcatagg
73441	gtgtcaagt	aggttttctt	gctatcttct	gttcccgggt	gogatggcag	aactggctga
73501	gccagatcag	ggtctgggtg	gcgtcagctg	attcatccag	tgagggtct	gcaagctatc
73561	tcaagcactg	atcttaaaatt	ttacaatagt	gatgttatcc	ccaggagcaa	ttttgggagg
73621	ttcagactct	tggagccaga	ggctgcata	cccctaaacc	gtaatttcta	atcttgtagc
73681	taatttgtaa	gtcttccaaa	ggccaactgg	tccccaggca	agaaacggat	attttcagga
73741	aagggtact	atcagttttg	tWtagagtc	aaaccatgaa	ctgaattcct	tcccaaagct
73801	agtttggcct	acaccaggga	atgaaaaagg	acagcttaaa	ggttagaagc	aagatggagt
73861	aggttaggct	ttatctcttt	cactgtcata	atttcctgag	ttataatttt	tgcaaagggt
73921	ggccaggcta	ccttctgcct	ctcatttctt	ggatattcca	tgaaaatttt	ccattcacaa
73981	aagcatttgc	taaaatatta	acagtactta	tttgcatggg	aaccaccaca	atgcacacag
74041	ctaggctgat	ccaggaaaaat	gttagatgtg	tttccccgat	tatcattttt	cacaaaaggt
74101	gattgatatt	atacctatct	ctttccctgt	gatttcagga	agtctcaaaa	tctcaaccag
74161	ccaagttagc	tgatcactcc	taccctgccc	ttataattct	aaaaaggtKg	attaaggtaa
74221	agaatattta	atttccttaa	catttaattt	tatccacatc	cttttaccac	ttgaggaaag
74281	tatggtggct	tgactgattt	aatgtgttgc	attaaacaac	tattctaattg	ataataagtc
74341	ttttattatg	aatctaacat	tctgtgcaca	gaatcagaca	caaaggactt	ttgaggaacc
74401	ttaattattca	aatggcttga	gagatattaa	aaaaataatt	acagtacaat	ggattgcata
74461	tgcatatcta	caaagagata	agtaaaaagg	aatggggaat	tgatgcact	aaatgtaca
74521	gaggacagac	atztatgcta	agagatcacc	aagcagagaa	ataatggaag	tcattccaag
74581	aaaaagaata	gcattattca	gagacctoga	gttgtatagc	acttgcccct	aaaatgtatg
74641	tctcagatct	tttccatctt	ctttaccag	ttcatcaggt	accagataa	gtgacttcac
74701	ctgctgtcct	tttcagagac	caaagggcag	gcttgctact	tacgtcagc	ttatgaaacc
74761	ctggcatgag	ttattcagat	tttaaaatag	ctattcctac	aaaaggcagt	aaattcatat
74821	tacttcttcc	attgtcttca	attgcaccat	aacattttga	attagaatct	gggaaggatt
74881	tctaattgagg	cactggaaaa	atacttgata	aatatttgag	aatttcttcc	tagtaattatt
74941	ttgagaaaaa	attgtaaaag	tgttgttgaa	cctaaaaact	agagcgtgaa	agcgtatgat
75001	ctcttgagag	tttatttatg	ctacaggggg	ttatgagctt	tatgtcaaac	gacttttgtt
75061	aaagaaagca	ttaggataag	gactcaaaaca	ttttcagtta	ttcttggtca	gactttaaca
75121	ttgtgtcagc	atgctttgtt	tactatgagc	atttaccat	tttatatgaa	ttatcacat
75181	ttccatttgc	catggatact	cctgttcctt	taccacaaag	tttaattact	gcttcttttt
75241	taagacaata	ttttctgaaa	acaataccgg	tctcattatt	taaWgtggaa	acatttctta
75301	gttttagcac	ttctacatgt	tgatttgc	ttttcatgac	ttataaatgc	ttagactatt
75361	actctataac	catgggaagc	tatttcaagt	gaggccagaa	aatgggaata	catattctca
75421	aaagaaaata	attcaacaag	atgttccatt	cattcttgct	ttggacaaaa	cattgttata
75481	ccatattgaca	aattaagaaa	cactaaatta	cagtcttgaa	aatacatgtt	tatctataaa
75541	atctctgact	taagttcaca	ttactttttt	ccatagcaca	gatttttaag	tagccttaat
75601	gatattgaagt	gaggcagaac	ttaggcattt	aggtgcttta	ttagaaagc	taagattcag
75661	ggaaatattt	ttaaacaact	tggaaatctaa	tcagtaaaaa	tatatataat	tttgaatcaa
75721	aaaatatatt	taaaaatata	ttgaatgcaa	cagtagtggt	gcattattct	taattgacat
75781	atatttgttt	catctaaatt	ttctcctttt	gacatcctaa	aagatgttgg	gcattctatat
75841	aaaaaattca	ttcaaagcct	tatcctcact	gtttttcaac	ctggtagtca	ctgtgagcta
75901	ccactttact	ccatttcatt	aattcatttc	ttaaactcatg	tcctagaact	taccaactag
75961	ggctactgct	atggtttgaa	tatgtcctcc	aatgttcatg	tgtggaaact	taattacat
76021	tttaattggt	ttagaaatg	ggggccttga	ataggtgatt	aggttatgag	ggctcttctc
76081	taagtggatt	aatgccatca	ttgcaggagt	gagttcttca	taaaaaggat	gattttgggt
76141	ttatttcact	ctgtcttgca	tactcacttc	tgcttctgc	ccttcttcca	tcttatgatc
76201	tttgccagg	gccaatgtca	ttcccttgga	cttcccagcc	ttcagaactg	tgaaaaataa
76261	atttcccttt	taaaataaatt	actctacctg	tggatttcca	ttatagtcca	gaaaatggac
76321	tgatgggtgc	tgtaatatat	taaaataata	caccaagtc	tgctattctc	aaaccattag
76381	ttctcatctt	tcattcttct	aggcccttta	ttccacaaat	tatcctgtag	aaatgttacg
76441	tcccctgact	ccacatttgc	ttccatattca	atgagctcct	ttctgactct	actttcttcc
76501	tgatgcatca	tttcaacaac	ttcttgccaa	tactctgaac	atgctcatct	atttgcctt
76561	ctgatgcaaa	agccctgaaa	atcttaagca	tttttcaaca	tacaaagcct	tacttttagat
76621	tgatatttcc	agaaaagaat	tgggtcataa	gttagagatt	atctgggttt	gggtgctttac
76681	agcttccctt	caaataaaac	tctgtccttc	aggtagtcaa	atataatgaat	gagttaacat
76741	ttgttagtgg	caacaagggt	catattcaat	atatttaaca	acagacaaa	tggaaataggc
76801	accatattaa	aaatgtcacc	tgttaaaagc	agccaatatc	agcttttatt	gaaagcaatg
76861	agaaataact	ccagaaagag	gtggggagac	ggaacaataa	aataaaatta	agggaggata
76921	cccaggtggc	taacagccaa	gaaaaatgca	ctgaatgaca	aagcttttga	atgaacaaga
76981	acaaagtcag	atctaggaat	attagcagca	aaagactgtc	aggcttcttt	ctgggtgctg

77041	ccataaaatg	acatagctat	aaccactttc	tgcctttgta	tattttctct	caagatcaaa
77101	attcttttagg	tggtagaata	tggttgctcc	aggctaggca	tgctcctccc	acccaaatat
77161	catgaaccaa	tcactatggg	cagttagatg	ggttccttcc	actggctcag	ccaagttcct
77221	aggccgcccc	agccatcaga	aaggtagagg	aatttctgga	agattgaaaa	gggataagtg
77281	gatgaaagca	agaggcacta	cagcatcttt	tgctgtcact	gaaaatttct	tttcttgatg
77341	gcttttMagg	agccctctcc	ttagtattat	agagagaaat	ctttgcttca	gttagagttc
77401	aagttaaact	acagacaaaa	aagctccttg	aaactctcct	tcaattactt	tcagagttga
77461	ccactttcca	ctggctattt	atctatctgt	ctgtcttcc	actcagcttg	gcgtgtaaat
77521	attgggtactt	cccagagctt	tgaccttact	ccttttttgt	ttcttactca	gcaaaccttt
77581	ccctacaaYc	tcatccacac	ccttggtctc	agcttccatg	tatatgctga	ctcctcaaac
77641	tgtagctcag	attaatcatt	tgcgtttttag	actcttaaat	gcaactgatc	tattgaacat
77701	ttccccttca	gataatgtgg	gggaatttcc	cacagaaaa	tcaaatttaa	tgtttccaaa
77761	taaaattcac	catccttctc	cccttccatg	tcattcctct	tctcctctgt	gtcatcgtag
77821	cctcattcat	ggcatcacca	tctgctgaaa	aatccaatct	gagagcttgg	aatcttcccta
77881	gactcaaccc	ctctctcaaa	ctcccttggtg	caatccacca	gcaaacgatg	gcagttctac
77941	tgactgctat	tgccccagta	aaagcactcg	tcatttctcc	taacatatct	ttctacctcc
78001	tgcgtaact	accatcaata	tgttttcaaa	atcctgctag	aaaaaaacta	aaaggcaaat
78061	ttgtctatgt	cccttaaaag	gctgcctaca	aaataRcatc	caaactcctt	ggctctgggaa
78121	gcaagatcat	ctgaagtttg	ttatctcccc	ctctatctag	tcccagctct	tgccatccct
78181	ataaaccaca	cttagaatgc	cagctgcacc	aaacagttct	gtttgtggaa	catggctggt
78241	catgtgcttc	ccatagacaaa	gacaaatcat	tctcactgag	gcttggtccaa	atatctgttc
78301	tgtcaataca	aatatatgtt	gttttacatt	acagtgggta	ttagtccagg	ttctctagag
78361	aaacagatct	ctagagaaaa	taagatacat	acatacatat	gagacaaaga	gaaagagaaa
78421	gtgagagaga	gagagactta	ttgtgaagta	agatactggg	tcacacaatt	atggaatctg
78481	agaaatccca	caatctgcc	tctgcaagct	ggaaactcag	gaaacccagt	gagagccaat
78541	ggggatatatt	ccagcctggg	tctgatggcc	tgagaactaa	gagtgctgag	tcagggaaaa
78601	gatcaRtaat	ccagctaaag	cactcaggca	gtgagcaaat	tcaacattct	ccatcttttt
78661	gttctataca	catcctcact	ggatgatgcc	ctctcacaat	gagcaaagcc	atctgcttta
78721	ctcagttcgc	taattcaaat	gctagtgtct	ttcccacaca	ccttcataga	cacactgaga
78781	aataatgttt	aacagatatt	ctgagtattc	catggcccag	tcaagttgac	atacaaaatt
78841	aagcatcaca	ctgagtttga	gKatatatat	tacacaacag	tatataactg	gacaggtgt
78901	ttttataatt	tacgattttc	ccatggtgaa	aatctccagc	attacctgga	gaatagattt
78961	ttaccttttt	acagaatgtt	acattttctc	ctttctggat	tattatgtgg	agtggtagaa
79021	aatatatgaa	atgtgagaac	agatctaagt	tcaaaaaat	tcatttccct	tattatcttt
79081	tatgaacttt	gataatttcc	tcttagtaaa	atatggatga	tatctacttt	gtacRtagct
79141	gtcgcRctta	attcagctaa	taggataaat	gctcctaaca	tctcagatgg	agtcagaggt
79201	tataaaccag	ggactatctt	tctccatgat	gattagcaat	ggaatgtcta	aaaagaaata
79261	agagaataaa	atatgagaca	taaatattta	aagagtgtac	attgaatagt	tggaacacctg
79321	agagaatcaa	gcaaaaaatt	attcaaaaca	gtaagaatgt	ttaagtggt	gctagctaca
79381	aaatataaat	ataagtcaac	aactgtcttg	tatttgaaga	ataatttaatt	tgaaaatatt
79441	actaaaaaag	acctcYatga	gaactaagta	taaaaatatcc	aagaacatgt	ttaaaaagaa
79501	atgtgtataa	ttattttaaga	aaacaatgaa	agtctgtcat	ataaaatttt	aatatgtggt
79561	gcccagcagt	gagtaattgt	tttgatgaat	atcagttccc	ttcctgcaRg	ctttaggcat
79621	tacagggtac	actgaaggca	acactacagc	ccttctgcac	agctacctta	agctttttct
79681	cataaagcag	atgaggcaga	tgtcatcaac	tgagagactg	gaaaaatgcca	acaattcttat
79741	ttgtagggaa	atgtcaataa	ataaaaaatt	aatgtggaaa	ccagttaaaa	ggcccatgaa
79801	agtcattttct	taccaggcaa	gatcaaatcc	tgagatcaac	tgctcacaat	tgcttagagt
79861	cggtgattag	aagagaaata	tataaagctt	agctttcaac	aaattaaact	ttcacatatt
79921	taaagacagt	tgaggcttac	ttttctgaaa	atgtgggggt	taagaaaaaa	aacaacccaa
79981	gcaaatagaa	aatggtagtt	cttttagactt	acagaagaaa	agcatgcacc	tgtttagaaa
80041	gtgttggtt	cctttttaat	tttttcaaa	gttttaggtt	ggtttRatga	catgatcttg
80101	atattattat	ggtaataaRc	tgcatttgct	aactatgttg	tatgagataa	ttattggttg
80161	ttttgaatct	aaagtctttg	agaaaaatat	ttaatcataa	gctaaattca	tcataattat
80221	gcacttttta	ggaggagtgt	tcataagata	gatttgaatt	tacattgagt	tttagattgg
80281	aatgataatg	ataagcttcc	ttaaaggtggc	tggaattttg	tcctcttcat	agatcaatca
80341	ctttttcac	tgcaagtaca	atttcatatt	tgttcaaaag	cttttgtaac	agagctctgt
80401	atttctcttg	caacgtatgt	tatctctttt	ttccctaaac	cctaattctt	agctggcctc
80461	actcccaatg	gaaaaaagat	gttttccagc	aaccgttgaa	ttcaatgtgg	acataattat
80521	aaggtcaatc	tgatgcccta	caaataaatg	tcttttggtt	aactttctag	aagagggttt
80581	atagcccttt	tgctcccttct	ttcctttctg	ctagctgaaa	aaaaattctt	tcacacatga
80641	cagagaaata	aactgttatt	ttattttaagc	cattgttatt	ttgagttact	gtatctcaag
80701	ataacttagt	tatctaagtt	gtcttagttt	actgtcactc	aagataacag	ttccaattac
80761	aacaaYattt	ataaaataaa	ttaaaaaaaa	caagaaggct	accttcacgt	taggtaaaaa
80821	aaaaattacc	caggagctaa	tgatttttta	aatcttgctg	agtgttttct	ggatcctatg

80881	atttcctggt	tgggtttatta	tctaaattgt	ttgcatttat	cagttactgc	tgctatgact
80941	caatcattat	aataacatga	tcgaaagcct	cctattttaat	atctactgca	atgcacattt
81001	gctgcactga	taatatcaa	gtctgccaaa	caggataaga	aaaacaaatg	ggatagaaga
81061	aactcataag	atatgctaag	cacagttaat	gaataattta	gcagcttgtr	aagaaaggaa
81121	aagctcataa	agatgtgtta	ggtagagggtg	gacaaataag	ttaattacct	aagaaacgtg
81181	tcagccatag	atacagtaac	taaaatgtgt	gtgttaagct	tatatagaaa	gggctggaag
81241	aaacaggata	aaagtaatat	tatgatagWt	gaagtcctga	aagggttgta	agcatgttat
81301	tcagccttct	agttgtctct	catttgctac	aactaagact	caatagctta	gttgaactga
81361	ctttttatgt	tactcttctg	agagaacaca	gtaaaaagct	caaacacaga	caaaagtagg
81421	catcctggaa	tcattctatg	cattgccagt	tagatgtctg	ggatgacctt	tgggaattac
81481	tccttttgga	gagcctcctt	catgtcttcc	aatacacacc	cctgttcctt	gggtgtttat
81541	actttggcta	tgcccagtg	tatactggac	actgtgtgtc	agttccacct	acttgctcaa
81601	agatcttact	tgtgaattct	gtcacacaaa	cactccaatt	ttcaacagct	acatctgtgc
81661	tctccctYac	catggctccc	tactacaaac	ttcaaacctc	cagaagataa	gtgtctgggg
81721	aactgactgg	tccttgagct	taaggaagca	ttgtagaatt	aatagcttgc	ttataataaa
81781	ggtggctgtc	accatcaaaa	gaacaatgaa	gtattatagc	tctggattca	ctactgagca
81841	aataacaaaa	tggattaaact	cttcagaaga	aaataacagt	gacagcctga	tatttcata
81901	tttcatcatg	gcatcaattg	aaattgataa	gggtctaaaa	ccaaatttat	caatagctgg
81961	gtagaagaga	ttaacatcct	agacattaat	gaattaaaaa	taaaagtcag	ctgctttaat
82021	aatccatgga	cagaaacaaa	aagagatttt	taaatggggc	caacactcag	aagagtgtga
82081	acaattgatta	atttttaaaa	acatacattt	gtgttttttt	ttcacttaga	agaaatctgc
82141	atttgtctag	ctctctaatt	ttttttaatt	attattatta	tactttaagt	tttagggtag
82201	ttgtgcacaa	tgtgcagggt	agttacatat	gtatacatgt	gccatgctgg	gggtgctgcac
82261	ccatcaactc	atcatttagc	attaggtata	tctcctaatt	ccatccctcc	ccccccccc
82321	cacccacaaa	cagtccccag	actgtgatgt	tcctcttctc	gtgtccatgt	gctctcattg
82381	ttcaattccc	acctgtgagt	gagaacatgc	gggtgcctggc	cctttgtccc	tgYgatagtt
82441	tactgagaat	gatgatttcc	aatttcatcc	atgtccctac	aaaggacatg	aactcatcat
82501	tttttatggc	tgcatagtat	tccatgggtg	atgtatgcca	catctttgac	aaacctgaSa
82561	aaaacaagca	atgggggaaag	gattccctat	ttataaaatg	gtgctgggaa	aactggctag
82621	ccatattgtag	aaagctgaaa	ctggatccct	tccttacacc	ttatacaaaa	attaattcaa
82681	gatggattaa	agacttaaac	gttagacctt	aaaccataaa	aactctagaa	gtaaactctag
82741	gcattaccat	tcaggacata	ggcatgggca	aggacttcat	gtctaaaaa	ccaaaagcaa
82801	tggcaacaaa	agccaaaatt	gacaaatggg	atctaattaa	actaaagagc	ttctgcacag
82861	caaaagaaac	taccatcaga	gtaatttttt	taaattcaga	aaattctatc	gtgtaaagga
82921	aaggggcggt	accggcatac	ctcagggttt	tgtcagaagc	tctgatttcc	ctgcctcttt
82981	attccctttc	tccaaactct	tctccacact	gattccagat	aactctttta	actggcaatc
83041	tgatttgtgt	actcctttgc	ttaaaattct	tcagttagtc	tttacaggat	aaaatctaaa
83101	gtcctttgce	tgatgaaaat	ccctttgcaa	tgcacactgt	ttactttcag	ggccttctct
83161	tccttctctc	tttccctgtg	tctctgtgtt	tcctgttaga	gtgcagtggt	aagtcagcct
83221	gctgggtctc	ccgctgggtg	cgactcccag	cagcacaaat	ggccttggat	ttcttccacc
83281	cccttccctt	tcaggcRtt	ttcactcagc	cctgcccaga	gaacagctgc	cactctgtcc
83341	aattctgtgg	gtctgtcagg	cctctgtgtc	tcccacttct	ccttctttct	cctggatccc
83401	gatcttttca	gcattcccct	cagtgtcctc	tttcacctgc	cacaaaaaaa	ataaaaaaaa
83461	aacaagccca	cattttaaac	aaattattag	agtcagatat	tttctccttt	cttaagacct
83521	agagagcttg	tcgattaaac	aaataaacaa	ataaaatatc	acacacagtc	cctcaatctc
83581	cattcccctg	ctcttggtga	gggtgatgag	caagacaata	tctgccttgt	aaatcagcaa
83641	gacggaaata	tacacataca	tatacatatt	catatattca	tacacatata	cctacatata
83701	tacttgtaca	tatgtatatg	tttctatgta	tatatgtgtg	tttgcatata	gaatatgtgt
83761	atcatgtttac	agaggtaaaa	actgttataa	aatgggaagt	tttaccttgg	ctcttaccag
83821	Rgagtgattt	tagggcaaga	tgcaactttc	tctctctctc	tcttcaaaat	ctttgtgcct
83881	ctctctctct	ttgtggctct	ctctcagtc	tttaattatt	tgctaagaat	tgtttttagca
83941	gtaaacaaaa	gacatgcaaa	attcttctct	tcataaaaat	tgtgttttaa	aaggagagga
84001	cagacaatac	acacaaataa	atctagaaag	ccagcagggt	aaagtgatgg	gcagctaggc
84061	agtgaccatg	gggaaggtaa	gggtgggtgat	gggtgtttgag	gtacatttct	caggcagctc
84121	Yccaaattgt	gattgaaagt	ttcctcYgat	gggtgggagaa	gcctaattgca	ttgtttgggt
84181	tcagtatagc	attttctttt	acataattat	tcataatgta	acttttttca	acattttgaa
84241	gcagcttacc	acttaagccc	catactgtca	tgtgaaacgt	tgaagcccat	taagatagac
84301	ctcaaacata	agaacaaaat	aagagaccag	acgcggtggc	tcacgcctgt	aatcccagca
84361	ctttggggag	ccaaggcggt	tggatcatga	ggtcaggagt	tcgagaccag	cctggccaaa
84421	tggtgaaacc	ccgtctctac	taaaaataga	aaaattagcc	aggtgcagtg	gcaggtgcct
84481	gtaatcccag	ctactcagga	ggctgaggca	ggagaatcac	ttgaaccogg	gagggttgca
84541	gtRagccaag	atcRtgccac	tgcactccag	cctgggcaac	gagcgagact	ccatctcaaa
84601	aaaaaaaaaa	aaaaaagaaa	gaagaagggt	gggctttgaa	ggaagatgct	gactctgaac
84661	cagcccacct	agactaagag	agtgtattgca	Rtgaacaaa	ggatccctaa	acagtcaagg

84721	caaaggRaca	gatgaaaggg	tttcgaggat	cttcataattg	aaggaacgca	agtctagcca
84781	acttagcatg	gttttttctta	gtaaaagcttt	tccataacca	agattttccc	tggttctttc
84841	atcttagaga	aattcataaa	ttggttggtt	tggggggaca	gttaaacaac	agaataaata
84901	atgttcaaaa	atcatctctt	tcatcaatct	tctgtaaagg	ctaaaggcat	aatacttaat
84961	tacaactgtg	tagaggcaat	acgatggctt	gctttctgga	attactgttt	ttctcattat
85021	ataaatatgt	attaagaaaa	gatggtatac	ttttaacttt	ctacttgga	ataattttga
85081	acatgcagaa	agttataaga	atagttcaaa	gaaaacaagg	atctgtttta	ctcagattca
85141	accattgtta	atattttgtc	catctgcttt	atctttgtta	tttgctatct	atctacattt
85201	ttttctgaac	tactaggaat	aaattacata	caacatgacc	ctttaactat	caatgcttaa
85261	gtgtgtattt	cctaagaatt	agaatattct	cttgcataac	tacattgttt	ttaacattta
85321	agaatctact	aacactattg	tcattttatc	aactgaacaa	agacctgtgt	agcttttgct
85381	ttgttttgtt	ttgttttgtt	tttcagttat	ttcagttatg	gatccactct	aggatcacac
85441	agtacattta	attttcaggc	ttcttgaatc	tttaactctg	aagagtttct	cagcttttac
85501	tttttataac	atttttgaag	aaattttcct	tccagttaat	aaaaggatca	tcacttgagg
85561	ttttcctaatt	gtttcctcag	aactagattc	atgttaacat	tcttagcagg	catattgatt
85621	tgcatactct	tagcaagagt	gtacacaaag	tcgttttatg	gtcttcttag	gatatacaYat
85681	gtggagacat	atcatgtttc	tctattttcc	gtgttaatat	aaattttgat	cacctgacca
85741	aggtacagtc	cagctactcc	agtgtgtagt	tattatttta	ttctcttgca	atagataagc
85801	aatccatattg	gagatacttc	aagactgtgc	aatatcctgc	tctacataaa	aattttcccc
85861	tgtgatattt	cactcatccg	tgattcttgc	atgaactgat	cttcattgtg	atgggtgcgc
85921	aatgatgact	tcataactcc	agcatgccct	gcacattcat	cagtcactac	tcaccttcta
85981	caatgaccaa	gagccctccc	ttctccccag	ttattttact	acatatttat	aatcagcatg
86041	tacttatgga	ttcctatttg	tgcaaagttt	ctgaatttat	tactatattt	acttattttg
86101	attctcaaac	ttaccaatgc	ttggcctgtg	agctcctgcg	tctttgtgac	ttatcccaga
86161	tggtatatac	tccttaattgt	aatgcattgtt	aggaccatga	aaaccaatgt	ttctcaacgt
86221	ctggtctctc	tttgtcctgc	atataccacc	cagggtgctt	aacaaatgca	gataacgtct
86281	tcgtactcaa	acctaactgga	ccagactttt	tggggatggg	caaactcaga	tgatttttat
86341	gaacaataaa	aatcaaagtc	gatgacatag	ataaggacag	agaaatctta	taccagatgg
86401	agttaccagg	caaagaacac	ttttccaaga	cgaatgcagt	aggggtcaaa	actattgtta
86461	caaaggatag	agattaaact	cagctctgct	gaaacaaaag	gaggcaggaa	ttttaaatgc
86521	tagggtgagc	ttatggaaaa	gtactgaagg	atttttaaga	ggaagttggt	caaagtgtatt
86581	agatggtttc	tattctctaa	ttggcactta	tcctagttag	gttcctatcc	tcccatagag
86641	actcagagat	gggccctatt	tttcttgatt	acgcttaaaa	ggaaggtgcc	cagatcctYR
86701	agaaagacat	tactgggttg	gaaaagattt	atatctcaaa	ggagcaaaac	aataattttac
86761	aatcagaagt	tttctagtaa	cttctctaaa	aRagggaact	taggggcccc	cagtcaggaa
86821	gaattctgtc	tRaagtttag	ccaaaactggg	tggaaatggt	aggctctttt	ggccagtcct
86881	cacaaaagca	gagactaaga	ctatcttggt	catctccgga	ttccctgtag	ctagctcagt
86941	ttctaccctg	ttgtaagctc	tcattaaata	catatgaaat	gaatgaatag	attaattgtt
87001	catttttatga	tcagagtcac	cagggaaatt	gtatttttat	tgcaRaaatt	tacttaaact
87061	ctgtaggtgt	ccctctgttt	cttttaggca	tattttctca	cctgttcaaa	ctctaaaaat
87121	ctcctggagg	gcttcataaaR	aatcaggtttc	cagagctgta	cctaagattt	cactaaatta
87181	aaatgtcagg	aagtagagag	gtaagtattct	gtattttta	gcaatttaaa	aatgctacca
87241	ggggagatga	tttaagcatc	tatttccctta	acatctcagc	accataggat	catattttta
87301	aagcctcaac	ctgtattttat	ttcaataaga	tgtcagggtga	ctctatttcag	ctcaaaatta
87361	agaMaggat	tcagttgcca	tttccaagtc	tgagggtgcc	tgtggtgta	ttgagcactg
87421	gtcttaatca	tacctcttct	tgtctgtagcc	tccggtttta	tgggaagcat	aagatcaagt
87481	cctgctgaga	ttttaaatgc	agcctccgct	actgctgtgt	tctgaaacct	agaggcgttg
87541	agctagtagc	atttgtatca	attactatta	aagtaaaagt	ttggcacaac	cacatctctt
87601	gagtaagcgt	gtaatgaaaa	aaaaaatcac	aaaatgcttt	tctctgaaca	cattctagat
87661	tctatcccca	ggatgggtggg	caaaaatttca	tcctcgagag	gctatggcac	atgtcccttg
87721	tttgtacccc	ttgaaaacta	gttggaagaa	ttgcctatat	aattattttca	actgggatta
87781	aaaagaaaaa	atacccgcta	atcagtatca	ataatataac	tcttacagaa	ttggtaagaa
87841	aacctctaaa	tcttgcatca	ctgtgaaaaat	ttaaagtttta	ttcaaaagct	gtagggatta
87901	taataagaat	aattttccact	gctgggttaag	agcttactaa	tttccaagca	tgcccttaag
87961	tagtccctat	gcatactctt	atttaataat	cataatggcc	acataaggta	gataaattat
88021	atgctcgttt	tacagataaa	gacactgagt	ctcaaagagc	tttaattagct	agctcaaggc
88081	catatttgcta	gtattacatt	agtgcaaaag	taactgtggt	atttgccatt	taaaaaaaa
88141	gtaatggcaa	aaaccacaat	tacttttgca	ccaatgtaat	acatgcggaa	gcagagcttc
88201	aaatgctagg	agctgtcttc	agagtttagca	ccaacgtaat	atttttaact	gggtggtctc
88261	aatcagagac	aattttgtcc	cagaagacat	tagcaaagtc	tgcaagcttt	ttgtttgta
88321	aagggagagt	tagagggtgt	gYgtgctatt	ggcattgtgt	gggtagaggc	catggatgct
88381	gctaaacacg	ctgtagtaca	cagaacagcc	tccacaacaa	gtaattatcc	agcccaacat
88441	atcaatagta	ccaacactta	aaaaacctgg	cacattggaa	aataaagcct	tttgaggNta
88501	ttttatgaag	tatacattat	atatacatat	atgtatatat	tatatagagt	ttcctgagaa

88561	ttgttcatgc	cggttaaagag	gaaaaagctRt	actgtcgcatt	tcaagctcta	cgtagtaata
88621	gaaacctgga	gagaacactg	aaaaactaaa	ttcctgtgga	acattaaaaa	tctttaattt
88681	agttaagtat	ttcatcaaac	agctttccag	aaccgttgat	tagaaagggtg	tctgtttcca
88741	cagagaatct	ctgagtcac	tctaccagat	gttttatctc	atgttttccc	caaagatatt
88801	tttatttcag	agcaagactc	attgaaatgt	taatttcggt	aaaattgagc	taacttaggt
88861	aacatactct	cagtgttata	ttccttttaga	gctaagattc	tatataattt	caattaagat
88921	gactcccctg	atttagaatt	aaataggtaa	aagtaagcaa	ttcgaatagg	gcaaaatctt
88981	aagatgttat	taaaatcaac	ataattttat	atttccttgc	ttgatttttt	tttgaagtca
89041	tgaattgaaa	tatatcttaa	ttggaaaagc	aataagagtg	aatgaggctc	tgctcgaaat
89101	taaatttaat	ctttcttttt	aatgatgcag	ccaacaaact	ttgggaaaaat	gtcaaagaaa
89161	tattttatgc	ttactttaaat	caatggacta	aatgaatact	tctttcttca	gtgtgatctt
89221	aaagtcaata	tacctaagcc	acaagcattt	aatatatatt	ataaaagagt	gattcccttt
89281	taaaccatca	tcgttagagt	ctttggaaaa	aatgacgata	aaataaaaatt	cttctgatag
89341	ctaaatcttt	ttgtgtggaa	cccaaataatt	ttaaaataata	ccatcgagct	tatcagttgc
89401	taacaataatt	cctgtcaaat	gtaaaatgct	gcatgtaaaa	tatttttcaa	attttaagtg
89461	gtaatgcatt	agcaatcaat	gtgcactttc	atttgtcaag	agtaagaagt	tattaccaag
89521	caaagcaatt	ttagagcaag	aagagcttta	aaagtgtgcc	cagtacatct	tggtcccttt
89581	agagctgaga	acactgatgt	cagagaggat	gacttttcca	aagtagtaga	gtatatggct
89641	aggactggct	gttaagccac	tgattccatg	cagcctcttc	tactacattg	tgccattatt
89701	cccaagaagt	aaaaattagt	tttttccctt	ttcaataact	agaatgccaa	gtatagtgt
89761	acgtaccaaa	ctagctgtta	tggtcgcag	atgccaaaga	gcttaaaatc	ttttgggaaa
89821	ggaataaatg	aaaaagataa	gaaatggtaa	tacaaggcac	aaactctggt	agagaagaga
89881	cattttccagt	tgggatgaaa	acaatcatct	ttaaggggaa	tggtttatct	gagtcagctc
89941	ttcaagtaga	aatgggattt	caagagaaag	agtaaaagga	caggagcttt	ctgagtgggg
90001	agactctatg	gaacaaatat	gcagacgtag	gaaattatag	gcatagctgg	gaaaaaacga
90061	aaagaacttg	tggtctgggg	tatgtgtatt	gcataaagaa	tgtgataata	aatttaggaa
90121	agataagttg	gaacctgggtg	actctgggtg	cttgcaatct	cctgggtgatt	gtagagagtc
90181	ctaattgcag	gccgggtgca	gtggctcaca	tctgtaatcc	tagcactttg	ggaggccaag
90241	gcgggtggat	cacctgaggt	caggagttcg	agatcagctc	ggccaacatg	ctgaaacccc
90301	gtctctacta	aaaatacaaa	aatcagctcg	gtgtgggtgg	acactgctta	atcccagcta
90361	ctcaggaggg	ctaggcaggg	gtatcgcttg	gaccggagg	gtggagggtg	gtacagtctg
90421	agattacacc	actgcactcc	agcctgggca	acagagcgag	cctccgcctc	aaaaaaaaaa
90481	aaaaagaaag	agtcctcatt	gcagctagaa	gtgaatgatg	agcagaattc	acaagaatgg
90541	ctgcttgaat	gtgaccctgt	caaaggctat	tttgtcacta	gagcctaagt	gaataatata
90601	taatatgaag	gaaatttttg	tttgtgtttc	taggatttat	tttatttttg	tggtcagcaa
90661	ttatgcaaac	caggtaacat	gatgaaactt	tgggaaagcta	ctaaaactgc	tgtacatgaa
90721	tggtgcaact	gaagttttta	aagttttctga	tgattcgtga	taccaaattc	attgtatttc
90781	tgacatttgt	aggtcagatt	ggaacttaca	aaagtaatga	attatggatg	tagttctgct
90841	taaagtgtgt	aaaacatttg	gagttttacca	gtcaagtttt	ttgtgataga	gaaaaaatga
90901	caattaaatg	gtctagaaca	taataaatta	attaaaaaaa	gaatttttaa	taaaaacttg
90961	aaagaatcta	tttgccattg	tttgactttt	tattcatggt	tttatgagaa	tccagggata
91021	ctattactgt	gaagtttagg	tgcttctctg	gaaagatttt	aagatttctt	cttcttacta
91081	gaaattattc	ttttaagag	atgctggcaa	ggagaatcat	gtgggtcagt	acacttttaa
91141	gccgattcta	ggatgcttga	ggcaagagta	caaacatgaa	aataaaatga	agatttttca
91201	agtttgaagg	aaatatctca	attggaaacc	gtggtttcct	atgacaaaa	aaaacatgga
91261	gttccaaaat	taatgatagc	taatgtttat	ttagcactgg	gaattaaatg	ttattctaag
91321	aggcttatat	gtactaagac	aaatacccaa	tcagggtatt	tatagagagg	gaaacggaag
91381	cacagcgaga	ttgagtagcc	tgccaatggt	gaagctcgaa	aataaatgta	ggcagctgta
91441	ctccagaaaa	cattctctta	actattgtgc	catgttgttt	cttaggacat	ggggagatat
91501	ccagctccag	gaaaaagtgg	aaagagcgtg	gaaaggacta	ttagattaat	taactctcct
91561	atcaaagtaa	ggtcaaaaaga	accctactct	tagtaaaagat	agaagtagat	taaaattcat
91621	aacattagaa	tattttttta	taattccttc	tgagaagaac	cagaactcag	taagagttagc
91681	tgactctact	tgaggcagaa	aatatcgagg	tcagcctgta	cattttgtta	ctgcaagaaa
91741	acatggatac	ttttaaagg	gtctgtggct	gaactaaaca	atttagaccc	aatgggcttt
91801	cgttgtctgt	actgtcaca	ctagagcccc	gggttatggt	gtgagagaat	ctaacaatga
91861	ctttaattat	tagaaccagt	tgacagccag	gcacggtggc	tcatgcttgt	aatcccagca
91921	ctttggggagg	ccgaggtggg	tggtatcaca	agtcaggagt	tcgagaccag	cctggccaac
91981	atagtgaaac	cccatctgta	ctaaaagtac	aaaaattagc	cagggtgtgat	ggcacgtgcc
92041	tgtagtccca	gtactcagg	aggctgaggt	gggagaatca	cttgaaacctg	tgaagcagag
92101	gttgacagaga	gcccaggcca	tgccattgca	ctccagcctg	gtgacagagt	gagactccat
92161	ctgaaaaaac	aaacaaaaca	acaaaaaaac	tgacatggag	aaaacctatt	aattgatgat
92221	gatgctcaat	atgagcaaac	aatatatagc	atgtctcatg	gactgcatgt	gtctccctaa
92281	agttcatatg	ttgaaacctg	aaactttcaa	gtaggtggca	acagggcctt	tggaagataa
92341	ttagagtcac	gagagtagag	ccctaattat	acatgggatt	agcgtcctta	taagaagagg

92401	catgagactg	atggtgtctt	tctccaccat	gtgaggataa	agcaagaata	tgtcaatctg
92461	caaccaggga	agaggacctt	caccaaatgt	cgaattagcg	ggcactttat	cttggactta
92521	ccagcctcca	gatctgttag	aaataaattt	ctgttgctta	agtcacatg	tccatgatgt
92581	ttttgttcta	gcagcttgaa	ctaagacagt	atcctagata	aaaatcatat	gtgagttatt
92641	ttatgggtac	aaaattgtca	acaacacaaa	tatatgacca	atggtgtctac	agaaatgacc
92701	acacagcctc	cccatacaaga	ttcagggtgtg	aactacagtg	aaattactgt	gtgaacatgg
92761	tccccgtatt	cagtatcaga	aacatcccta	ataagagaat	taaatcaaag	caggagactg
92821	gctgcaggca	aatagatttt	ctatgaggat	atgacctgt	tgcttcatgt	agaaagcaac
92881	atcttgata	tgtgtaaata	tacttgatta	aaagtgtagc	ttgtttgctt	ttttcctaag
92941	ttctctctga	ctactgatat	ataccctgat	gacaaaaggc	ctgggcatgg	caaaaatgga
93001	tttctcatat	ccaaaagacc	ccttccttgc	tcaaacagca	tcacaagcac	tatccgcacc
93061	tattagatat	tgtgcaggcc	aagaaggcgg	tgctgtgata	catgcatctg	cttcttctcc
93121	aaggggtcaa	cccataccca	atactttttt	tgcatcttgg	tttaattaaa	ttggaaatgt
93181	aatcaaatat	gtcaatgagt	atattttcag	tatatctttt	cccatttcta	ggattatctg
93241	aaataatata	acctttaata	taagttttcg	aagttctgtt	agacttatat	atcggtcata
93301	ttactcaacc	ttttgtacca	tttatttgat	tgatgggtaa	ttgtattctg	atttatgttt
93361	tacaataata	ccaaatttat	aggattgttg	gagatgatgc	ctagcacatc	tgaagggtccc
93421	caaaaaatgt	agtcttttat	tactactact	tcaaagtga	caaagagagt	caactttata
93481	ctoctagagc	cccttaagaa	tatagcataa	tggtcgccaa	catgaacatt	gtgggcagca
93541	cagtgtaccc	ctgctcacag	tccttaggaa	tccttgtagt	ccagtgtgct	agagtttctc
93601	cgacactgga	aaactgttgc	ctttattcaa	tttgtctcag	tagctctcaa	cctcatgaag
93661	aaaaagtcga	tggttaggaga	gaagaaagag	atcaaagaga	agcatttggg	gcagagatta
93721	cggaacgat	gtaactactg	caagatggtc	agagaagggt	ttatatattg	aaataatgtt
93781	caaataattga	tcatgaccca	gaacttttcc	ttttgttctg	ggatacagtg	agatgaagga
93841	gagtacattt	tgcatttctca	tttgcccttc	ctttgtatat	tcttttgtaa	cacctatcat
93901	taaatgaaac	tccaccggt	actcaagcca	ccccaatgtg	gtatcagtc	tttaatttat
93961	aatacacatg	tgctaaaaag	tctgagcttt	tccgaaacct	aaaatctctc	tcagtccatg
94021	ttctgaaagg	acagtaagaa	aaggcacttt	acaactatta	ctgaacatct	aggtgggttag
94081	agagaccaga	caaatacacag	aagcactgtg	cacacccata	gtaaagggaa	ttaggaagaa
94141	caggtgactg	cagcttttca	tatttaaagg	ttgtctggaa	atcagaaaca	gcatatttga
94201	ctactctagt	aaactccaaa	gtatatagtg	actcagacag	ggtaagttaa	ttctcatata
94261	tttaaagtgc	aaatagatag	ttttattctg	ccagtagctc	ttctccaata	agtgattcag
94321	ggacttgggc	ccgttgattc	atgtgtctct	gtgatcttca	agtccttatt	tcccagggtta
94381	ccaagtttgt	ctgtctcaat	cctgaagaag	aaagagcatg	gtgggtctag	cacccatggg
94441	caatatgtta	cattcaagcc	tgagaatagc	acatctcact	ttcactcaca	gcccattggc
94501	aatcatattg	taactgcaaa	ggatactggc	aaatgcagtc	tagctgtgtg	cccagaaaga
94561	agagtaagta	gatttagtta	actgttagcc	acacacctgt	gggaaaaatg	ctaataaaaa
94621	atcatacttt	aaaattttta	ctgagaataa	cacacctcac	tttactcac	agtccattgg
94681	caagtattca	gtaactgcaa	aggatactgg	gaaatgtagt	ctggctgtgt	gcccagaacg
94741	aaagtaagt	agatttagtc	aactgttagc	cacacacctg	ctggaaaaat	gctaataaaa
94801	aatcatactt	taaaattttta	actgagaata	gcacacctca	ctttactca	cagttccattg
94861	gtaagtattc	agtaactgag	aaggatactg	ggaaatgtag	tctggctgtg	tgcccagaat
94921	gaagagtaag	tagatttagt	caactgttag	ccacacatct	gttggaataa	cgctaataaa
94981	aaaatcatat	tttaaaattt	taactgtctc	tttgagaagc	attttgctt	ttaaaattat
95041	ataaaatata	actgtattca	tattttatat	acaggtatat	atttataatt	ctcacaataa
95101	gttcatattg	aaagagtcaa	tctaagaatt	catcactgac	cccacctggg	cttgaataaa
95161	aattcattac	tagctggaca	atcctatcag	atatcattta	tttgacagg	ctctgtacag
95221	taattctcta	gtattccagt	gtactgtttg	gtataactta	tagtataagc	agagtggctc
95281	ttacttgtta	tcatcttggt	catttcatta	gaacttaact	aatagatgga	ttgtgtctgt
95341	aatcccagca	ctctgggggg	ctgagggtgg	aggattactt	gagaccagga	gtttgagacc
95401	atcctgggca	acatagggag	accctgtctc	taaaaaataa	aaataaatta	gccaggcatg
95461	gtgtgcgaca	cctgtagtcc	cagctactct	gtaaagctga	aRcaggaggg	tcacttgagc
95521	ctaggatgtc	gaggctgcag	tgagggtatga	tcataccatt	gcactccagc	ctgagcaaga
95581	gagcaagaac	cccatactctg	aaaaataaaa	aaaaaagaac	ttaactaata	gacattagca
95641	aggcccccac	tacagactgt	tcacacctat	gtttataact	gacactaaag	ctgcaacaag
95701	ggctgcattg	acagaattcc	agggtgttag	aattgcaaaag	tgaattgtctg	ttttaattaa
95761	tgcttgaaag	taaattgcag	ttttcatgca	tataatacat	tattattaac	tatagtcacc
95821	atgctgtcct	ttaggtctcc	aaaacttatt	tatcttataa	tcaaatattt	gtaccttttg
95881	actaatatct	ctccatttct	tccacctga	ccctgggttaa	ccacaattct	actttatttc
95941	ctgtgttttt	ttttagattc	cacataaag	tgagataagg	aggattttgt	ctttccgtgt
96001	tgtgcttatt	tcacttgaaa	ttgtttaaga	gagaagattt	taagtattat	aagcacacag
96061	taaaaggggg	gtaattttct	gaagtgtatg	ctatgttgat	agtcatttca	taatgtatac
96121	atatatcaaa	acatcatgtt	gtacaccttg	aatatataca	atttttgttt	gtcaattata
96181	cctcaataaaa	gttgataaaa	gatcaattgt	attaacagaa	catacattaa	aaagacaaaa

```

96241   tcatccatca caataatggt ttccatatatt tcatagtatg cacagtctaa acaaggacct
96301   tcaacagatt ttaatTTTTgt atttgttgggt accattcaaaa cagttttcat ctgggtatta
96361   taataccaag tatgcatgta gtgctcactg gaaaaaaaaa tgaatgattt ttgatccact
96421   gtcattacaa cttttttcat gaagtaactg gaaaattgca tatattatcc tgtcaccccc
96481   aatTTtatag tcctgatgcc ttttataaat ctatggctaa gggagacctg cagaatcaga
96541   aaattccctc ggagtgggtc tgggcttggg aacaatctaa gtagcctagg gaagctatta
96601   ctgaaagtgg tttctcccca gctgagaaat gtttcaaatt taaagattta agtctccca
96661   accaattcac atgacatgtc agatgagtgt gttaaagagc tcaaaacaat ttctttaaaa
96721   ttctaacttc atctctatag ggaaagaaat aagttgggat gttgtttctt aattggcaca
96781   ctcattgagt gtctctccat gaagggcaat cagttagttt ctatgaagca aagaccagta
96841   agctctggcc tctgcattaa caagtgtca tttagttaaga gtcacaggaa agcagaatta
96901   aggataagga caaaatctca ggtggaatg tgggggaaag gtgtgctgag taatttcttt
96961   ctcatgtYat agagaaagat gagcaagaat atagacaaag atcgaatagg aatgtgtctg
97021   cctagactgt gttagactat aaaaattagc ttcattgata cgtgcttaaa agcaggagta
97081   ggattcaaaa tatTTaccaa ctggtatggt ttgggcacca acccttcagt acagctcctg
97141   agtctaagca accgatcggg caatgagtac cagcagtggg cgccctactg aagaggccaa
97201   ttgagcactc agtgtggcaa gagaagggtg gaggaaaaca ggggctactg catggacaga
97261   ctatTTgatg agaaaacaac actcaaagat gatttctgtc tttgctgtgc atggccaacc
97321   ctgatgtctc atcaggagat gcaggccagg tgaaataaag ccactcaggg ctattgagat
97381   cctgtgatgc tgtgatttat cagagtcagc cagaggactc atgtcatctc ttggtgtata
97441   tccaaaaYgc tggcgTTTgc tggactacta gactgctaga atgctgacga aatgattgta
97501   atggcttccct ccctaacttc tctccaatct ctaattgtct gtattttttc cataatgaat
97561   atcatcaYct gaacttgcat tatttgttca ctttctccc tcaactaaaa tataagcttc
97621   atgagagcag agctttgttt tattattggt gttgttccact gctgtatccc tagtgcctaa
97681   gaccatgtct ggtacatagt agactctcaa taaatactgg tgtatttatg aataagtaat
97741   aacagtaact gtcacctagt aagggttaa tctatttcag aagtaggctc ttttaattgca
97801   ttacctgatt cttttagtagc ttaatgtcct tataagtTTt tgcacagtaa tggagtgaca
97861   ttaacaatag aaatctaaat atcccctgcc tccactcttg gatttccata ggggtgcctcc
97921   catctcagca caaactgacc gcaagctcaa atcacctga ggcaaaactt tctaattttt
97981   ttgttccctt tgtctctttt tcatatgggt agcccccgtt gccctgtaag ccataggaat
98041   gtgcttccca ctgccacagt ggaagtgtt atcaaaaaca aaattggcac aggacaaagt
98101   tcaacaaaga aatTTTTatt caaagtatt gcactaaggg agagaggcca gaattaagtc
98161   tgaacttaac tgaactgagc tgaagggaa aactcaagct atctatgttt gctaattggc
98221   aatattcaaa gagggaaaac tcaaggtgtc cgtgtttgtt aattggcttt acccaaagga
98281   aaagtaaatg tgcacttatc ttcatgtcaa tagacagttt tacaatgtgg agcaagatac
98341   ccacaaaagc taggcattctg tctcaccaca gagactgagc gagagggtt ctatcttcc
98401   tgatgatagg atagccatca aagggtggc tcccaggtcc ttaagaaaga tactcgtagg
98461   ttataaacgc tggccagaag cttttaaatt gatataatc tcaaaggagc agagaaggaa
98521   tttgaaatta caggctttct aaagtaaatg ttctaagaaa agcgagggtca gggccctgga
98581   gtgtctaaag tttcatcatg ctgaaggctg cttctgtcag tgttcaccat gggaacaccc
98641   agtgcaggcg tgtctcttgc tactctcaga gccaaagggtc attaagatgc cacagcagca
98701   ctgtggtcca tgaaaaagga gaaYgccctt ctctcccaaa ggggtaccatg cggtaacttg
98761   aacctttaag tagcactttc tgtgtaattc tagttcccat ctctctgtcc actggatcca
98821   atatcaaact aaaggactag aacatggcac agagtatggc atcctctccc tctccctcat
98881   ccaacttacc tttccctcta actgcttaat tgtcaatcct gaaaactagt tttctctact
98941   tttagcaaaa

```

SNW1 genomic sequence (SEQ ID NO: 3)

>14:76161051-76258200

```

1       aactcaagtg atcctcctgt gttggcctcc caaagcaatg ggattagagg tatatgccac
61      tgtgcctggc caaagcaagg agtattaact ctgggggttaa tcttagctgt actgcttact
121     tgcctcgggc atttttagatg agtcatTTaa tgtcttagag cctcagtttc tcatctatat
181     agtaacagta gagtcaccca aaattaaggc aaaaatggWaa taagaaaact taagcaggct
241     aatTTTggct tcaaaaacta ctttttcccta tgacatcata ctttccaac ttcttatggg
301     gctgaatcat gaaagggtgtc tgggtgactg aaatatcttg agtcacaaaag ctatttactt
361     cagttagggc taacagtaat gcttttgaga cacttgataa gagttggtat ttaaaccagag
421     ctataaggaa caccaagctg gacagggtga gaaaatgaaa agagaaggcc ggtcgcagt
481     gctcacgcct gtaatcccag cactttggga ggctgaggca ggcggatcac gaggtcagga
541     gtccgagacc aacctggcca acatggtgaa acccgtctc tactaaaaat acaaaaatta
601     gttgggtgta gtggcgggag cctgaaatcc cagctacgag aggctggggc accagaattg

```

661	cttcaacctg	ggagggtgaaa	gttgagctga	gctgagattg	tgccactgta	ctccagccta
721	ggaacagagc	aagactccgt	ctaaagaaaa	aaaaaaaaaa	aaaaaagaga	atgaaaagag
781	agtgcgtgtaa	acagacagaa	aaaatgttgc	agtgaagacac	tgggaatcct	gtctgctaaa
841	agtaaaagct	aaaaaatcct	gtttgccctc	ttctagatga	aatgtccaac	tgacgctaca
901	acttctgctt	ctgttaatta	cttagaaatg	aatttagtat	ttcaggggct	gaagttttct
961	ttgtatgtga	gtatctaata	gatggctcat	tacattgggt	aagagacca	ctatgcagtc
1021	attgaagtca	aaacattaaa	ggaaaaaacc	tttattcttc	cctgaatagt	aaacctaggg
1081	tctgccttgc	cttaaaatat	atgaaacagg	ccgagcgggtg	gctcacacct	gtaattccaa
1141	cactttggga	agccaaggca	ggtagatcac	ctgaggtcag	gagtttgggt	ccagactggc
1201	caacatggcg	aaaccccgctc	tctactaaaa	atacaaaaat	tagccaggca	tgggtggcgcg
1261	tgccgtgtagt	cccagctact	tggagggctg	aagcagaaga	atcatttgaa	cctaggagat
1321	ggagggtgca	ctgagccaag	atcacgccac	tgcaactcag	cctgggtgac	acagtgaagc
1381	gctgtctcag	taaaaaaaag	tatatatat	atatatacac	acacatatat	acacacacaY
1441	aYatatatat	atatgaaaca	ttctctttga	aagagaatga	tcttttatct	tatcaacttc
1501	tccaagtgtg	gagattcctg	gacaaaatac	caaaagtgtg	ctgatgtata	accagtctct
1561	ataataaaaag	ggctatatga	aacctctga	cagtactatc	agctgagttt	tccaggtacc
1621	taggcactgg	taaaacactt	ttaagcattt	tgctcatgatt	tctttaatag	gccatactta
1681	atactgcact	ataattcaac	acactctcag	agggtgatgtt	agtggacctt	ttctatcata
1741	tgctaaactg	aacgcacacca	ctgagctttt	ataaaaagggt	ttctctttcc	tcccttgcca
1801	gactgttaac	caccactttc	cccaccagtc	tcactcagctg	gcttgtaata	gtaacttctc
1861	ttacttacta	ctttaatctg	ctttgagctg	aggaactaaa	cacccaagct	ttcatcagtc
1921	caacacatac	tgtacttagt	actgttttag	gtcctgtact	agacacataa	tggctcattc
1981	cctgatggct	gacattctaa	aaagcatgta	catattatgg	taaaatgtag	gtaaaaggag
2041	ctgaatgaga	gtctgaagcg	ctctggagag	aaaaagtata	ttattattgg	ggtgtggact
2101	ttctgtattt	tgacaactct	aagaccactt	ttttcccta	acaaatgaag	gtgctgaaaa
2161	gctgagctca	agtatgcaat	cttaaacact	taatactgta	gcataccaat	tttatcagag
2221	ctgttaggaa	gcttggaag	aatgcctcat	ctagacatct	gtggacaaca	ccaagagcaa
2281	attaagggac	tgcatatcct	caacgttaaa	gtggattgta	aagaaagaaa	atataaattg
2341	gactacagtt	ttgggtcccc	gttttccctc	ttgcctgtta	tgttccctta	aaaaaataat
2401	tttttaaaaa	tctattttact	tgagaacagg	ttatgagact	ggctaatttt	ttgtattttt
2461	ggcggaacg	ggggctcttg	catgcccag	gctggtctcg	aacccctggg	ctcaggtcag
2521	ccacccgcct	cgccctccga	agtgcgggat	tacaggcttg	agccaccgct	cccggactgt
2581	gttccccttc	ttaatctcca	tggggaaaag	Ytgagagaga	gaacggtgcc	tgtattccat
2641	atatacttgt	ttgccattta	tttcaaacac	ggagggattc	tcagaatcct	tcaatgacct
2701	ccttctaaaa	gagccagagg	agaagacagg	gactagtttt	acgcacagtg	gcttaacaaa
2761	cagcgccgca	tcctgttcag	ttctgggatt	cctcgggagt	ggtttttttt	aagctgcaga
2821	agtctaccct	accgtctaca	aagacgttct	atggggtttt	acaaaaaccc	aaacagaaat
2881	gtgaggagac	agttacagta	agaaataaac	acggattccc	aggctctttt	gactgagcgc
2941	ccccgccagg	aatggccgcc	taaagaaaca	ggcatctcaa	cgctttattg	ctaagtgtga
3001	cggttagtgg	ggagttcgac	tctgcagcca	aaggagggtca	ggaatcggtc	aataacaatc
3061	ggctctggag	tgaaaagagg	cgaagaagac	acgtaaagtc	ctccaagacg	cggggcaaaag
3121	cgatggagag	acgcgcgcga	ctcctctctc	tgtacctttt	gggcaccagg	acccttgcca
3181	cgggctgcgt	gggcccgcga	gaagtogatg	accccttcca	ggtctgcggt	cccgggcccgg
3241	ctctgacggt	agaagcggaa	aagtttccga	aaggcgctct	ccccgggctc	agtcgccaga
3301	gtcgccacag	agccccaggc	cgctgccatc	ttccccatct	cgcgccctat	accctctgat
3361	ccggaagcag	attctctcgt	gttccggatc	cggaattttt	ttccggggcc	gcgacctcgg
3421	ctcgagaagg	tgcttttagtc	tgaagatggc	ggcctcagca	gcgagagggtg	ctgcccgcgt
3481	gcgtagaagt	atcaatcagc	cggttgcttt	tgtgagaaga	attccttgga	ctgcccgcgt
3541	gagtgaagtga	tggagatgtg	ggagttagtg	aatttttagg	tccaagtgat	cctcaaagct
3601	tctgcttttt	gcaggggtact	taagttagcg	tgggtggctga	attaggagac	tgctcctgag
3661	catgcaagtg	agcagaaaaga	aattttgttt	gcagttaacc	gtttagggat	gtcccttgag
3721	cttttgggggt	cacgaatgca	ccggctcttg	cctttaggag	tgatttgggtg	taaagattta
3781	ttacttggta	accaagattt	acaatttaga	agggataatt	ttgactgtag	tgtggaggat
3841	ggattcaagg	tggcaagatt	gtaggcggtg	agaggagcta	aaaggcaatt	tcattaattt
3901	acttaaaacg	tgaggacctg	aattagggcg	atggtagaag	gatggtgggtc	agatgcatta
3961	atgttttaaa	agcaaattag	atgccgtttc	gtgatggcct	aagtgggaga	agtgatcgga
4021	ataaaggaaa	gggagaaatg	atttctgggt	ttttagcatt	ggtgaattgt	ttgttttagaa
4081	gttaggagca	ggccaagaga	agatgagtaa	agtttgaatg	tgttgacttt	gagaaacttg
4141	tgagggtttcc	gtcaggcaat	tggacaaaagg	ttgtatatca	ttgtggctga	gaagtggact
4201	ccgaagctag	accgttgggg	ttagaattcc	aacttggcca	tttaatcggt	gttatcactt
4261	gaataactga	agttagtatg	ggaaaagggc	tcacaactgg	gcctcagttt	actcatMgat
4321	aaaatggagt	caataatacc	caactcatag	gattgttgag	aggattgagt	ttatatgcaa
4381	acttcttaga	cttgtgttag	atacagtgta	agtactcagc	aaatgatagt	tattttcacg
4441	gtaactaaaa	ttgaaactta	cgggagaggc	ctagtctaca	gataaagggtt	ggaataatta


```

4501 gttctgaaat ctttcaccac cttcaatttt tgtcagctgc agttttgaga tgagggtttt
4561 atatttgctt accaacaatgt aaacattcgt ggtcactcct tttgtagata ttcattgtttc
4621 aggtcaaccc ctggaaacaa aatggagcaa gatttagttt cttctaagtt ttttagtggct
4681 cagatgtgag tagtagaatg ggaaataggt tggtttgtgc cccRtgtttt aaacatgtgg
4741 ttggtgaaga ccaagcccag taaagtctcc cgtagtccct gtgaaacKtt atggaaaggc
4801 tgaggagttt cattagcatc ttgcggatgc aagtttccag acaataaagg gctagttaa
4861 aacctttaa aagagtcagc attatgtaca gcttcgaaac agtttctcct tgttgttcca
4921 ttgattcagc aaatatattt ttgagtatct attctgtggt aggcattgtt cttggtgctt
4981 gcagctgttt ttccctcacc attttaggca atttaggaag agggttacct tccagagctg
5041 gtttggtttt aaatcagtgt cttctaggac ttttaattatc tggataaaaa catatgttta
5101 acattatttt gttctccagt atcatattct gtgagacctt gatcttgcca gttggtttga
5161 cttaaccatt agagcccttc aaaatttagt ctctgtggt tKttggaagg attagtgga
5221 ttttacttat tcatttagcc aatattgagg ccctgctgtt tattggtcac tgtaaacata
5281 gtggtttata tgctgagggg gcttacaatt agtggtggg atgaagagac ttagaacaaa
5341 tataggcaat tgttgaacca gtgaggtttt ggtgaactag gagaactaaa tagatattat
5401 aaaaagtaaa tcagtgaagt acaagcactg tagttcagaa gagaacaga tcgttagtt
5461 tttagtcttt aagattatat actctaaatc tttgaagtct cacagggtg gtacattgtc
5521 ttgtactgtg ataggagccc aataaagtaa tactcagtgg atgtacttgg tttgggtctt
5581 gaaggatagt tatagtgtta caagttagcc agggggagat aggtattacca acatattggg
5641 gggagagaac aacatatgca aagagggtaa cccgcaggac agataaagat aattcaatta
5701 gttgtactag agagtaacta ggggttagca gtgggaaata aggcaggaag ggtaggtgag
5761 ggcaactcaa cacattttatY taaattgtga aagaccagta agttgggatt ttatttggaa
5821 atcttcagta taccgggaaa cacactggat tgtaatttgt ctttttcttc atttgtatct
5881 gctcacaggg atcatagtcc acatagcaca gtaactggtg tgaataaact cgatatttgc
5941 tgaggatgga gactgcagcc aactcaacaa cttYtctgtc catctcacca ccaagtctgg
6001 tacattattg ctttatagtc agtaccatag ctgcttccta aatcttaagc ctttaaaggaa
6061 atgcaaagta gttcctgtcc acaagaaatc tggtgacct gtctttctac agtctaagat
6121 agagaatcca aaatatagta actactttta atgttttctt ctgccacaca ggtcagctga
6181 aagaacactt tgcacagttc ggccatgtca gaagtgcat tttacctttt gtaagtatta
6241 aggaaaagta ggtgggaggt ggggatggct aatgggtaca aaaaaaatag aatgaataag
6301 acctatttga tagtacaaca ggtgtactat ggtcaataat aattgtacat tttaaaataa
6361 ctacagagtgt aattggatta tttgttaact aaaggataaa tgcctgaggg gatggatacc
6421 ccataagtca aaggataaat gcttgagggg atagataccc catgctccat gatgtgctta
6481 atttacattg catgcctata tcaaaacatc tcatgtactc cataataaat atacatattt
6541 atatatataa atatataaat atatacacct aataatataa aaataatataa aaatatatac
6601 ccaaataaat atatacacct actatacact cacaaaaatt gttSaaaaaa agccttaaag
6661 tgttttaaaa aattcttatc ttgcaacagg gtctcgctct gtccccctgc aggggtgcag
6721 tggcctgatc tcactgcagc ctctgcctcc caggcttaag tgatcttctt gccttggcct
6781 cctgagtagc tgggactata gacatgtgcc accatgccc a gtaactttt ttggtatttt
6841 ttagtggaga cagggtttcg ccatactgtc cagactggaa aattttttta attttcattt
6901 tctttaagag acagggcttt gttctgtcac cccaggtgga gtgcagtgt gcagtcattg
6961 ctcactgcaa cctctgcctc ccaggcttag gtgatcctcc cacctcagcc tctgagtag
7021 ctgggactac aggcattgcac caccacacct gggtaatttt tgtatttttt gtagagatgg
7081 ggtttcgcca tgttgcttat gctggtctcg tcgaactcct gggctcaagc tatccgcaca
7141 cctcggcctc ccaaaatggt gagattacag gtgtgagcca ctgtgctggc ttatttttgt
7201 ggagatgggg gtcttgccat gttgccagg ctgatcttga actcctggct tcaagcaatc
7261 ctctgcatt ggccctccaa agtgcaggga ttacaggcgt gagccactgt gccggccta
7321 cagtctttat tgattacatt ttgagtata cttgattaaa aagtgaataa gctttatcat
7381 aggtatttgt gttagttttt ttattgttgt ataacaaat accacaagcR taatggctca
7441 aaaacaacac atctctcaca gctgagcaca gtttagctgg gtccctctct taggatctca
7501 catagctata tattccaggt gtcagctggg gctgtggtgt gtctcatcaa aggtcaact
7561 ggggaaaaag actgacttgc aagctccctc atgtggttgg caaaattcct tttttttttt
7621 ttttgagatg gagtctttct cccttgccca agctggagtt gcagtgtgct agtctcagct
7681 tactgcaacc tctgcctcct ggttcaagc cattctcctg cctcagcctc ctgagtagct
7741 gggactacag gcatgcgcca ccacgacggg ctaatttttt tttttttttt ttttttgaga
7801 cggagtctca ctctgtggcc tagctggag tgcaagtggc cgatctcgcc tcaccacaag
7861 ctctgcctcc tgggttcacg ccattctcct gcctcagcct cctgagtagc tgggactaca
7921 ggcccccgc actacgcYcg gctaattttt tgtattttta gtagagatgg agtttcgcta
7981 tgttgccag gctggtctcg aactcctgac ctcaagtga ccccccgcct gggcttccca
8041 aaatgctggg attacaggtg tgagccttca tgcccggcca taattcatct ttaaYttag
8101 gactgaggtc tgttttcttg ttgacctttg gctctgagtt gggagctgct ctcacctagg
8161 aactccctgc agttccttgt taMctagttt atctgtaggc tgtcttaca taaggcagct
8221 tgcttcttta aagccagcaa gggcttctcc agtctgctta agatggcatc ttatacagaa
8281 cataatcaca ggagtgacat cccattactt tYgctgtatt cttttggtta gaagcaagtc

```


8341	agggtttctcc	tMcacccaag	gagggggagga	gattatatag	gggtgaacat	ggggaatgca
8401	tggagatcat	ggggcccatc	ttagaattct	gcctactaca	gtatttcta	aaggacttga
8461	gattttcta	aattggcaga	ggatatttta	agttttatct	tcattgagta	ccattttctt
8521	tttttagattt	ataactaata	aaacaataat	tttatattaa	ttacatttcc	tttgctttct
8581	tgtgagggtt	tagaaatcag	cctataaatt	tatcctgggc	tgtatttttc	tgtgggcttt
8641	tttaggggtg	gatgtgagta	catagagttc	acacacactt	caactgaaac	cttttttctt
8701	tgcagcactg	taagggtgtg	ggtaaaataa	caacataaag	gaacttaaat	tcttaattct
8761	tggtttaagt	ttacttgact	cccaaagact	aacagattag	ttatatatag	tgccggtgaa
8821	gggtgtaggga	atctggcact	caagttgttg	agaagttcaa	ctggtagcct	cagttctgtt
8881	gtgtagttgg	caaaataacca	aagttaagaa	tttatatttg	ctttgatcta	gcggttactt
8941	ttctaagaat	ttatcctaca	gaaatgtgtg	gtcaagaata	ttccaaagca	gcattattta
9001	atgtcaagta	agaagaagaa	acttaaattg	tcactcactg	gattgattag	atattaagtc
9061	actgatgcaa	tggaaatttta	tatagttgaa	tagaaaagtt	gggcacatgt	attatggaca
9121	tgaatatattt	tctgtaagtt	tgaaaacttc	aattttttca	agagcattta	aataaaaagt
9181	tccttgcaaa	aaaatgaata	gattttttatt	gggtccataaa	taaaatagga	cctgcaaaga
9241	atagatcttta	gatttaaata	aaaggcggt	ttgatgttgg	taaaacaaaa	taaaatatga
9301	caaaaaaaat	atttttttgt	tccattcact	gtactctagt	tggaaatYct	ggttcctttt
9361	ttgagacagg	ttctctcNct	cactcaggct	ggagtgcagt	ggcacaaatca	agttcactgc
9421	aggcttgaac	tcctgggttc	aagggtacct	tctgcctctg	cctcctgagt	agctgggatt
9481	agtcgtgtgc	caccacacct	gactaaattt	ttgtactctt	tgtagaaatg	gggcttcacc
9541	atattggcca	ggctgggtct	gagctcccg	actcaagcag	tctactcacc	tcagctccca
9601	aagtgtgtgt	attacaRgca	tgagccacca	cacctggact	gagttatttt	ttttgagaca
9661	gccttgctct	gtcggccagg	ctggagtga	gaggcgtgat	ctcagctcac	tgcaatcgct
9721	gcctccYggg	ttcaagtgat	cccccttct	cagtcaccca	gatagctggg	actacaggca
9781	tgggccacca	cacccagcta	atttttgtat	ttttagtaga	gatggagttt	cacYatgttg
9841	gccaggctgg	tcccaaactc	ctgacctcaa	gccatccatc	cacctcagcc	tcccaaaKg
9901	ctgggattga	gccactgcac	ctggcctatg	ttttaatata	aacaaatttc	ttaatgcctt
9961	tgactgtgga	agtggtcaca	tgacaataaa	ggaatgcttt	agaaattgtt	gtatctacaa
10021	tgaatggatc	attgggacat	caccagaaac	tttccgtctt	ctaggttaat	ccttataaag
10081	ttatccttga	agtagaggag	caattgtta	attttgtttt	ttttggcaaa	gcacagcatg
10141	acttaaatga	gactaatgag	gacctgtttg	tcactctacag	ggactagaca	tcttaatatc
10201	tgttttctgt	Ygaagggctg	gaggggcagg	gtacatccaa	gcactctaga	atctgttttc
10261	cgtatgcatg	gtatagggtg	aaacttttgc	acaagggaaa	atagcaaaat	tgtttaaatt
10321	aaccagaagt	ctgtaactag	gttgagtaag	ggctaaatat	tatcttttct	tgttttNcat
10381	taggaacacg	tgatgatagt	cgtgggtggg	gtgagggtga	ctggattcag	aaatgagtgc
10441	aacaaggctt	tgaggcttaa	ttttctagtt	tctagaaacg	aagcttagaa	caagagtact
10501	gatcaaaatt	tcccttagct	attactgtta	gtatggtttg	catctagaat	gacacagaaa
10561	gctcatgacc	cgtatcttct	ctaattggcag	caaggaaaac	aggaagctta	acaacacagg
10621	gatttagcct	tgggtaatca	ggaaggggtg	aaagcaagg	tactgacact	tataaatgtc
10681	tgtgatccga	ttttYggctg	ggcatgggtg	cttatgcctg	taattttcaSY	actttgggag
10741	gccagatggg	gaagattgct	tgagccagg	agtttgagac	cggcctgggt	aagatgggaa
10801	gaccccatct	ctgcaaaaaa	tataaaaaat	gaggcagaca	tgggtggtatg	tacctgtgat
10861	cccagctact	gaagaggctg	aggtgggagg	attgcttgaa	cccaggaagt	tgaggctgca
10921	gtgagttgcg	ttcataccac	tgacccccag	cctgggtgac	agagcaagac	cctgcctaaa
10981	aataaaataa	aattaaatta	aaaagtccga	tttttgccag	atgaaaaagt	tggtagtttg
11041	atttggaact	catggaaact	ttctgaagtt	catgattaat	cttttcttat	attttgaatt
11101	tttaggacaa	ggagactggc	tttcacagag	gtttgggttg	ggttcagttt	tcttcagaag
11161	aaggacttgc	gaatgcacta	caacaggaaa	atcatattat	agatggagta	aaggtaaatt
11221	tattttctatg	ccagatacat	acatgatata	catgtaggta	ctatttaatt	atgtatcaat
11281	taaatagatc	ataaatgtgg	aatgatgggg	acttgagag	atttgtaact	gaagctgcct
11341	taatgcttgt	taagaatggt	gggtggagg	ctgggcacgg	tggctcacgc	ctgtaatccc
11401	agcacttttg	gaggccgagg	cgggcggatc	acRaggtcag	gagatcgaga	ccatcctggc
11461	taacacgggt	aaaccccgtc	tctactaaaa	atacaaaaaa	ttagcctggc	aaggtggcgg
11521	gcgcctgtag	tcccagctac	tcgggagggt	gaggcaggag	aatggcgtga	acccagggg
11581	tgggagcctg	cagtgaagccg	agattgcgcc	actgcactcc	agcctgggcg	atagcgagac
11641	tccatctcaa	acaaacaaac	aaacaaaaaa	acaacaaaaa	aagaatggtg	cgggtgtagg
11701	ccaggcgag	tggttcatgc	ctataatccc	agcacttttg	gaggccgagg	cgggggggat
11761	caccagaggt	caggagtttg	agaccagcct	gaccaacatg	gagaaacccc	atctctacta
11821	aaaatacaaa	aaaaattaga	tgggcgtggg	ggcacatgcc	tgtagtccca	gctactcagg
11881	agactgaggc	aggagaaccc	aagagacaga	ggttgagtg	aactaaaaatc	gcgccattgc
11941	actccagcct	Sggcaacaag	agcaaaactc	catctcaaaa	aaaaaaaaaaa	aaaaaaaatg
12001	gggggtggta	gagtggccag	ttttaaaagg	tccaggaaag	tacaagatga	ccctttaatg
12061	ggtagcaaga	tatgtaacac	agtggacata	acagtctact	tttctttttt	ctttttgaga
12121	gtcttgctct	tgttgccag	gctagagtgc	aatggtgcaa	tcttgatca	ctgcaacctc

12181	cgccctcctgg	gttcaagcga	ttctcctgcc	tcaacctccc	gagtagctgg	gattacaggc
12241	atgcgccacc	acgcactgct	aatttttttt	gtatttttag	tagagattgg	gtttcatcat
12301	gttgggtcagg	ctgggtctcaa	actcctgacc	tcaggcaatc	cgccacctc	ggcctcccaa
12361	agtgcctggga	ttacaggcgt	gagccacagc	gcctggccag	cagtctgctt	ttcagggtcag
12421	gcaagctggg	gctgaaatct	tggcttacc	accatgtgat	cctgggcca	actgatttac
12481	tattttatatt	taaatttttag	ttttaatgag	acagggcctt	gctatgttgc	agggtgggtg
12541	tgaactcctg	gcctcaagca	atcctcttgc	ctcagcttgc	caaaatgctg	gcattacagg
12601	cctgagccac	tgcccctggg	ctgatttaca	atttttaatt	gaggggttag	gagtcataat
12661	accaacaaag	atgggtaaaag	gaatttgagg	agaaaaYgaa	agaacaagg	acaaataaaa
12721	actgcctggg	ttattcatta	ttcatactga	ctccctaata	aagtgtttt	cagtttgaat
12781	gtttttgcag	acattgcctt	tcctccctta	cagtgccttat	ttttttaagg	ttcagggttca
12841	cactagaagg	ccaaaaacttc	cgcaaacatc	tgatgatgaa	aagaaagatt	tttgagactg
12901	cagcctatta	ataaagttaa	cataactgag	aattttgtct	aaatgtttt	atttgaacaa
12961	aatagttgca	ccaagcaaga	gKttactttg	cccactccaa	attaaaacag	agcacaatag
13021	gggcaaaatt	tattttggcag	gacagttcca	gtatgtgaac	atcttctctc	tcactgtggt
13081	tggttgtaact	ttactcatat	gcagctgttc	ttacacaaaa	cattaacccc	aaactactag
13141	tgctacataa	aagtaagtgg	tcttgacttt	gtatgtgggg	cagcatgttc	tataaatgct
13201	gaaaggtggg	agaagcacia	acacacccca	ctctttaaaa	aaaactaaat	aattcaaagt
13261	agaattttct	atccccccca	tttctccagt	aataaaaagt	agtgtcggga	tctggcacc
13321	agattttgggt	tttatcctga	ccattttaca	agtgttcccc	catatgactt	gcactcttag
13381	ggttatgggt	aagagttcat	tcattttgga	gagacctgtg	cctattcctt	cctcctcttc
13441	ttgccttcat	gctcgtgttc	cttggggcgg	ctgctatctg	agggtctttt	agagccacca
13501	tgctgtttgg	cttcttccaa	aaacttgtcc	aaaccaaag	gatcttcctc	aaactgcact
13561	ggctcttctc	ggcctctctg	tctacggctc	gaaccagaaa	actccttgct	gggaacaaat
13621	ctaaggaaaa	ggagaaaaaa	ctatgaacta	ctttgctttc	accagtgtaa	acactgaaat
13681	tgagttgtat	ggccttggc	ctgttggtct	ttattctggc	ttctaggtca	tcaccataca
13741	tgctcttgtc	cagattttta	ctgggcttat	aaatactctg	ggccatatct	ttaccacctc
13801	ttcaggcttg	atcataaaca	ttataaattt	catcttctcc	acctgcaa	ccactgtcca
13861	taccctgtgg	aaaaatggat	gacattaaat	aaaagtgtaa	acattgttta	tcaaaagctg
13921	aatcataaca	tgtttacagt	aaacccttgc	taatgtacag	ctcacatttt	caacaactctg
13981	aactgcaaat	taaaaagtct	tatttaaaaa	actggtagca	catgaacttc	ataacaaact
14041	tggtattcttt	tttttttttt	tttcagacag	agtctcactc	tgtaacccag	gctggagtgc
14101	agtgggtgtga	tcttggctca	ctgcaaccac	cgctctccag	ttccagtgat	cctcctgect
14161	ccgcctccca	ggtagctggg	attacaggca	tgtgccacca	caccagcta	atttttgtat
14221	tttttagtaga	gacagagttt	caccatgttg	gccaggtggg	tectgaacc	ctgacctcag
14281	gtgatgcct	gccttgggtc	cacaagttgc	tacatagctc	actggagctg	tgacctccct
14341	gtgctcaggt	gttccctcca	cctcagcctc	ccaagtagct	gggagtacag	gtgtgtgcca
14401	ccatgtccag	ctaattttYt	gtagacaggg	ttttgccata	ttgccaggc	cggtcttaaa
14461	ctgggttcaa	gcaatctgcc	tgcttaggtc	tcacaaagt	Ytaggattac	aggcgtgagc
14521	caactgccc	agccagaaat	attttcatca	caatttttaa	atgttttaag	cttaagcttc
14581	taacaggtga	ttatcttgaa	aatattttact	tacgtgaaat	ttcaccctta	gatccctttg
14641	ctcagctggg	aagttgccat	tgtatagtgg	ttaagaaagc	tttgaaattg	gctgggcaca
14701	gtggctcacg	cctgtaatcc	cagcactttg	ggaggccaag	gcaggaggat	cacgaggtca
14761	ggagttcaag	accgccttga	cccaacgtga	tgaaccctg	tctctactaa	ggatacaaaa
14821	attagcctgg	catggtggca	catgcctgtg	gataatccca	gctactcagg	aggctgaggc
14881	aggagaatcc	cttgaacctg	ggaggtggag	gttgcaagtga	gccgagatgg	caccactgca
14941	ctccagcctg	ggcgacagag	tgggactcca	tatcagaaaa	caMaaacaaa	caaaaaaaac
15001	tttgaaatca	agacacccag	tttgaatctt	aattccatca	gtcacaaagt	aaacctgtgt
15061	ttcaattttc	tcatctataa	cataagagat	aatagtatta	tttcatagga	tttttgaag
15121	aatttagata	ataacaagac	ttagaacaat	gccaaagaata	taaaatgcat	ctaataaatt
15181	atcacttaga	agagcttact	taaaacctcc	aagtcataat	taagtttccc	actctacttt
15241	ccagtagaac	ttacacttct	gtattcttgt	gtgtgcgcaY	gcgcccgcgt	gtgttttccc
15301	aggctttaaag	tgctcctagt	agctaattgg	aattggtact	agaaatcaga	tttcttgaca
15361	gccagcctag	tacttttaatc	ccgtttttaca	gcttcatttt	taaaagaatc	ccaaattgat
15421	atatttctat	agcattaaaa	gaaaaatcca	aaaatcaaca	acaaaaaaac	ccaaRaggca
15481	cttttgtagg	acactgactc	ttttattttt	ttgagataga	gtctcagtca	cccaggctag
15541	agtgcagtg	cacaatcttg	gctcaccaag	caattctcgt	gcctcagcct	cccaagtagt
15601	ggggattaca	ggcgtgcacc	accacacctg	gctagttttc	gtatttttag	tagagatggg
15661	gtttcaccat	gttggccagg	ctggttgtga	actcctgacc	tcagggtgat	tgcccgtctc
15721	ggcctccaaa	tgtgttggga	ttacaggcaa	aagccaccat	gcctggcctc	ctggttcta
15781	ttttagaaac	aaaaggttga	cataccacaa	caccaagttt	ttaatcacct	aatttctttc
15841	ttctcctagt	gtctcaaatc	tcagcccact	ataagaagtt	aaaatcaRaa	ctgtttctatc
15901	ttgtctcaag	tttcaacata	caacttccaa	tttttttgct	atcatgctgg	catttacatc
15961	caaatgtaaa	gtatgWtaat	tttcgttttc	ctgaggatag	gtattcacat	caacacacac

16021	aatatgctgt	tcctaaactt	gtaccttgga	ttgggtgaag	agccttttgt	catactgaac
16081	ttcattggaa	gtccgaggat	taggaacacc	gagagcaata	acttcactga	tatcccgatt
16141	ttcattttctc	tgaagtttcg	acctattttg	aaatacgaca	tcactaactg	aaaactcttt
16201	atagacaagc	tgaatatgtt	caatagaaat	acaaccaatc	tcagaaacaa	aggatgtgaa
16261	tatgtcacat	tcacatcttg	tgaacatata	ttccttgatg	atacatMatc	aagctggcag
16321	cccttataaaa	gcattttcaga	ttagagggag	agtttttaca	ttaggacgtc	ctgtagctc
16381	aaaagaacat	ctaacagatt	cgttgacatg	gaattcttta	atataagcaa	tcaccttaca
16441	aMcatgctga	gatataaaat	tgctatgggc	tatatattact	atcattatat	gatgggttgg
16501	gggttgggga	agtggataca	atttagatgt	agcacacgaa	acatttttaa	taagctagaa
16561	acagctcttt	tatctaacca	catttagtaa	ttttatccca	ctagaatttt	gaggaaaaca
16621	tgtcaatgac	tgaattagta	tgatttttga	cattcagatt	aatttgggtt	cctggctggt
16681	gcattatata	tgcaacctac	atcccaatga	aacctaggta	actgtcatta	cttgttttac
16741	taaaaactca	ggtcacaaaa	tgatttctat	aggctgtttg	attttcatcc	agttacaatt
16801	tcacagcttt	tctatgttgt	caattctgat	agccacttta	ctcgttttga	gggcccatt
16861	tatccacaga	gccattgtaa	cttttacaac	ttaggaatta	ctaaccaaga	ctgtcattaa
16921	ttaaacactt	tcacactttg	ctcctccttc	tacatgtact	ttaccctcaa	ctctcaagag
16981	taagagtctt	tgcttctcca	tgacatcatt	tatccaggga	gatgaaatat	gccaattaat
17041	accagcagga	atctttttca	tgcttattta	tttttatttt	tttatttttt	attttttatg
17101	acacggagtc	tcactctgtc	gccagggtcg	gaatgcagtg	gcacgatctt	ggctccctga
17161	aacctccgcc	tccagggttc	aagtgaattt	cccacttcgg	cctctcagat	agctgggact
17221	acggcatgt	ggcaccacac	ccagcttaatt	tttgYatttt	tagtagagat	ggggtttcac
17281	catgttggtc	aggctgggtc	tgaactcctg	acctcaagtg	attcgcgtgc	tttggcctgc
17341	caaagtgtcg	ggattacagg	tgtagccac	tgtagccggc	ctcttttttg	tgtttaaaat
17401	atgacttgaa	aaatgtatta	ggctcatggg	atgatcagag	cacagaaaac	tgttcatttg
17461	ggaagaaaat	attctatatt	tgcatttaac	cctcactcag	aattccacct	tagtattgca
17521	gcgttccctc	tacttttcag	aactttttag	cctcagcctt	tggtgagcca	taaaactccc
17581	caattcctac	tttactctat	gccctcatat	atatatttag	tttgttctat	acgtttatac
17641	ttactaagag	tgctttgcaa	attttctcaa	attattttat	ggatagctct	gtctttctga
17701	actgtaagtt	aaactactcc	taaaaatgag	ctgattttct	acttctttca	aaactgctct
17761	tctcctcaat	tgtactagat	gcagtactat	gcataataaa	tgatctcaat	aaaagcatga
17821	aagcttattt	cctgtctaaa	gaggcacata	atccatttat	ctcaccattt	ctcagtagaa
17881	aagaacggac	tataacagtc	caaattttact	actacatcct	tttcaaaaat	ctgatctcga
17941	ttccgttttc	tatattccca	tcattctcaa	ttctgactta	agtgatggaa	gattaattcc
18001	aaagacagtc	aactagagct	aagacaaaag	agttaagtac	aagaaaatta	agtgctaatt
18061	tattaagtta	tcaaactaaa	gaagaaaggg	agagaagcaa	aactgatctt	agtggaactaa
18121	aactctgtct	taaaaggtag	aaaaaggcca	ctcttcatga	ttggtagata	tctgtttttt
18181	tttttttttt	ttttgtgaga	tgtagtctcg	ctctgtcgcc	caggctggaa	tgtagtggcg
18241	tgatctcggc	tcactgcaag	ctccacctcc	cgggttcacg	ccattctcct	gcctcagcct
18301	gccaagtatg	tgtagctaca	ggcgctcgcc	acgacctatg	cgggctaatt	ttttgtactt
18361	ttagtagaca	cagggtttca	tcgtgttagc	caggatgggc	tcgatctcct	gacctcgtga
18421	tccgcccga	ttggcctccc	aaagtgtctg	gattacaggg	gtgaaccact	gcgcccggcc
18481	accatgattg	gtacacatct	taaataaacac	ctacctctta	tcaggagctg	ccctggaaag
18541	attccgggtc	tgctgtctct	cttttcgcct	gtcatgccgg	atttcatccc	tctcacgtgc
18601	ctccccatcc	tctgtgtgaa	cagttgaaaa	gtacttcact	gtaatatgta	ttatatgtgt
18661	gcatacgtgt	acacgtgtgt	atataataaa	ttttttaaaa	agagacagca	tctcactctg
18721	tcacccaggc	tggagtgtag	tagcatgaac	atggttcact	gcagccacta	cctcctggat
18781	tcaagtgatc	ctcctgcctc	atcctcccga	gtacctggga	ccacagatgt	gtgccaccat
18841	gcctggctaa	tttttaaaat	ttttgtaaag	atgggttccc	acttatgttg	tcgaggctgg
18901	tctcaaaact	caggctctag	tgatcttccc	acctaggcct	cccaaagKgc	tggagtccaa
18961	ggtgtgagcc	actgcactca	gccccattag	tattttgttt	gttttgagac	ggagtttcgc
19021	tcttgctgcc	caggctggaa	tgcatgggag	cgatctcagc	tcaccgcaac	ctctgcctcc
19081	cgggttcaag	cgattctcct	gcctcagcct	cccaagtagc	tggtgattaca	gggtgtgacc
19141	accatgccgg	gctaattttt	tatttttagt	agagatggag	tttcaccaag	ctagtctcaa
19201	actcctgacc	tcaagtgtat	tgccgcaccc	agccaatatt	ctgaacaaga	gaatattttt
19261	taaccttaga	aaattccaaa	cttctttgta	gtcacactca	gcctttttaa	ccttacttgc
19321	aattaatttta	agtctaataa	atgatatatt	tttaattttt	gaatgtgaat	tctagaatca
19381	agcatgcggc	taaaatattt	aagggtaaaa	tgtattgctg	tcttcaactt	acttWgaaat
19441	gcaatgcagt	aaatcggaca	gacagatgga	cagaaagatg	ggcaaatatg	tgataaagca
19501	aatttatcaa	tgtaatttgc	tttttaaaat	tatttttgat	atataatcat	tatacatatt
19561	tatggggatg	atatatttga	tacatgaaca	aaatgtgtaa	taatacaaat	agggttaatta
19621	ggatatccat	cacctcaaac	atttggttgt	gttgggaaca	ttctaaatct	tctagctatt
19681	ttgcaatata	caataaatta	tcattaacta	ttcttctctt	ctagctattt	tgcaatatata
19741	aataaatttat	caYtaactat	tcacccacct	gtactaggaa	cattagaact	tactctttct
19801	aactgtattt	ctgtaccatc	aaccaatctc	tcttcattcc	ctgtccccaa	catccttccc

19861	aaccactgggt	aaccaccatt	ctactctctg	cttccttgag	accaactttt	ttagctccca
19921	catacaagtg	ggaagattca	gtattctggt	ttctgtgccg	ggcttattca	cttaataacc
19981	ttcagttcca	tccatattgt	tgcaaattac	atgatttcat	YMttttttat	ggctgagtaa
20041	tatcccatgt	tgtatataga	ccacattttc	tttatccatt	caccactggg	acacaacttg
20101	attccgtatc	ttggtaactg	taaagagcac	tgcaataaac	atggaagtgc	agatatccct
20161	ttttgatttg	atatactgat	ttccttttgg	atatataccc	agcagtgaat	tgctggatca
20221	cgtagtggct	ctattttcag	ttttttgaag	aactgccata	ctgttttcca	tagtggctgt
20281	actaatttat	attcttacta	aaaatgtact	agtgttcccc	cccttctgta	tcctcgccag
20341	catctattat	tgcccttttta	atgatagcca	ttttaactgg	agtgcataaa	tatctcactg
20401	tggttttgat	ttgaatttcc	ctgatgatta	gtgatgctga	acattttttc	atacacctgt
20461	tggccataag	tatgtcttct	tttgagaccg	tctattcaga	tcRtttgccc	attttaaaat
20521	cagattatct	cgtttctttg	cgattgagtt	cttgtacatt	ctgggttatta	attctttctt
20581	aaatggataK	tttgcaacca	ttttcttcca	ttctgtagg	ggcctcttca	gtctgttgac
20641	tgctgtgcag	aagcttttca	gcttgacgca	atccccattg	tctatatttg	cttttgctgt
20701	ctgtgctttt	gagattttac	ccaaaaaatc	tttgtccaga	ccaatgtcct	gaagcatgct
20761	cccagtgttt	tcacttagca	gtttcatagt	tcaggcctta	catttaagtc	tttaaatccat
20821	tttgatttga	ttcttgtatg	tgataagaaa	taagaggctc	attttcattc	ttttgcatgt
20881	gactattcgg	ttttcacagc	attatgtatt	gaagagacta	tcctttcccc	aatgtatgct
20941	ttgggcatca	gcgttaaaaa	tccacgcatt	ggctgggcat	gggtgggtgt	catgcctgta
21001	acgcactttg	ggaggctgag	gcgggtggat	cacttgaggc	caggagggtg	agaccaccct
21061	gggtgtaatg	gcaaaaccct	gtctcataaa	taacacaaaa	attagccagg	cggtgggtg
21121	cggtgccttag	tcccagttac	ttgggaggcc	aagggtgggag	gatagcttga	gtctgagagg
21181	cggagggtgc	agtgagctaa	gattacacca	ctgcactcca	ggctggatga	cagagtgaga
21241	cgttgtctcc	aaacaatcaa	aaattccatg	catttatagc	cgatccattt	atttctgggt
21301	tctctattct	gttccactgg	tctatgtgtc	tgtttttatg	gcagtaccat	gctgctttgg
21361	ttactatagc	cttgtagtat	attttgaagc	caggcagtgt	gatgcctcca	gttttttttc
21421	tttttattca	ggattgcttt	ggttatttgg	ggtcttttgt	ggttccatac	aaattttagt
21481	attgtttttt	ctatttcttt	gaagaatgtc	attggtattt	tgatagggat	tgcatgtgtg
21541	ctagagctag	ctttgggtag	tatgaacatt	ttacaataaa	taatttttcc	aatcatgag
21601	catgggctat	ctttccattt	ttagagtgtc	ttctacaatt	tctgttttta	aaatttaatt
21661	ttttttcttt	taaaaagaga	tagggcttta	ctgtgttgcc	cagggtggtc	ttgaacacct
21721	gacctcaagt	aatccttctg	ccttggcctc	ccaaagtgtt	gggattacag	gtgtgagcca
21781	ccatgcccag	cctcttcaat	tacKttcatc	agtcttttat	cattttcttt	gtggcgatct
21841	ttcacttctt	tggttaagtt	tattcctagg	aggctatgtg	cagtgaactca	tgccgtgaat
21901	accagcactt	tgggaggcYg	aggcaggcag	atcacctgag	gtctgagacc	agcctggcca
21961	acgtggtgaa	accccatctc	tcctaaaaaW	aaaaaaaatt	agccaggcgt	gggtgtgtgc
22021	gcctgttgct	ctagctactc	aggaggcgga	gacacaagga	ttgcttgagc	ctgggaagcg
22081	gaggttgcag	tgtgctgaga	tcatgccact	gcactccagc	ctgggtgaca	agagtgagac
22141	tccgtctcaa	aaaaaaaaaa	cctttattcc	taggtagtgt	ttggtagcta	ttataaatgg
22201	gactgccttc	ttgtttttca	gcttgggtat	cctatttgta	tatagaaatg	ctacagattt
22261	ttgtatgttg	gttttgtatc	ttgtaatgtt	actgaattcg	tttatcagtt	tttaagttttt
22321	tggtggcatc	attaggtttt	tctaaatata	ggatcatact	gtctctgaca	aagtataatt
22381	tgatgtcttg	tttttttttg	tttttgagat	gRagtcttgc	tttgttgccc	aggctggagt
22441	gcagtggcat	gatctcagct	cgctgtaacc	tctgcctcct	gggctcaagc	aattcttgtg
22501	actcagcgtc	ctgagtagct	gggattacag	gtgaaYgcc	tcataaccag	ctaatttttg
22561	tatttttagt	acagacgggg	tttcatcatg	ttggccaggc	tggtcttgaa	ctcctgacct
22621	caagtgtattg	cccattctcag	cctcccaaag	tgctgggatt	acagggtgtga	gccactgtgc
22681	ctggccttcc	tttccaatct	ggatgccttt	tatttctttc	tattgcctaa	atttctctgg
22741	caaggacttc	cagtcccatg	ttgagttaca	gtggtgaaag	tgggcatcct	tgcccttctc
22801	cagatctctg	agtaaaagact	tccaattgtt	ctccattcag	tatgtcagct	gtgggtttgt
22861	catatatggc	cttaagtatt	tttagttatg	ttccttctaa	gcccactttg	ttgagagttt
22921	ttatcatgaa	gggatgttga	attttttcaa	atgctttttc	agtttctact	gaaatgatgc
22981	tatggttctt	attcttggtt	ctattaatgt	gatgtatcac	atttactgat	ttgcatatac
23041	tgaaccatct	ttgcatcctt	gggatgaagg	atagttcagt	tgattatgat	gaatgatatt
23101	ttttaatgtg	ttgaaatctg	tttcttagaa	ttttgttaag	gatttttgca	tctaagatta
23161	tcagggtatg	tagcctgtgg	ttttcttttt	ttgtgtatat	ttgtctagtt	ttggtctcag
23221	tataactaatg	accttgtaga	atgaatttgg	aagcattccc	ttctcttcaa	ttttttggaa
23281	tagtctgagt	agaactggta	tttagttcct	taaatgtttg	gaaaaattcg	gcagtgaagc
23341	tgtcaaatcc	tgagcttttc	tttcatggga	gcaatgttaa	ttgttaaaaa	aaaaaaaaaa
23401	attagtgtag	aatctaggtg	gtaggagata	cRggtgttca	tcataataatt	ctttcagctt
23461	ttctgtAKgt	ttgaaaatgt	tcataataaa	atgtcaggaa	aggctgggtg	tggtgcctca
23521	tgccgtgtaat	cccagcactt	tgagaggctg	aggcagggtg	atcacctgag	gtcaggagtt
23581	tgagaccagc	ctggccaata	tggtaaaacc	ccatctctac	taataataca	aaaaatttagc
23641	tgggcatggt	gatgggcacc	tttaatccca	gctacttgga	aggcttaggc	aggagaattg

23701	cttgaaccca	ggaggcggag	gttgagtcga	gctgagattg	catcattaca	cttcagcctg
23761	ggcaacaaga	gYgaaactcc	gtttcaaaaa	aaccaaataa	taataataat	aaaaaagtca
23821	gaaaaaagga	attatgcata	atgtttccct	tatatattca	aagattctat	tgaataaaga
23881	actgtcagct	tattttccctt	aggggctggc	tggatcttga	tttccactgc	atttatatct
23941	atcttcattt	ttttctaccc	tcgccagctc	gaggggaatg	agtacagtta	agagcataag
24001	atttggaagca	agcctgtctg	gattcaaatc	ctgtttctat	agccaattaa	ccatgtgacc
24061	tgcgcagatt	acttaatact	ataagcctta	gtttcctcct	ctgtaaaaca	gtcacataaa
24121	tagtatcact	acatagggtt	atcccaggga	ttatgagata	atgcaaataa	agtactcaga
24181	gtgggtgccc	ccataagaag	cacctgataa	atgttagttt	ttactactat	ctggcaagag
24241	attcattaaa	atttttaatt	tcttgagaac	gattctgttg	tctttttctt	attagtttag
24301	agaactcaaa	tgcagtaagt	tttaYttcta	attatcttct	atggtttagt	atacaaccag
24361	gcagggtat	atatctcacc	ttattaagtc	attcatatat	aactgaggga	agtggatcac
24421	tacttaaaac	cctgggttatt	tccctatggt	tgaactctag	actcttgtaa	ccaattacct
24481	ggcacctcca	tcttaaacctt	aacatatact	taactataat	aaattcttca	tctttccctc
24541	ccaaacctgt	tccctcctgca	gtcttcccaa	tattagtaaa	taaaagtaat	atttactttt
24601	atcctgagca	attggcaatt	taatccttcc	caactgctta	ggataaaaag	tctggagtca
24661	tccttaactc	ctactttttt	tcgtatctcc	acctaatacc	caaacaaatt	ttgtcaactc
24721	tactaccaaa	atacatccaa	tatgggacca	tgtttttggt	acctcttggt	acYatctctc
24781	ttacagttta	tttggaaggc	attagcccat	gccatttcca	cagtattcta	tttttaacac
24841	agcagccagg	gaggtctttt	aaaaatata	gttaggggaa	atgggtcata	ggtacaaagt
24901	ctcatttatg	caggatgagt	aagtcttgga	gatctaatat	actgtatggt	aactatagtt
24961	aacagtactg	catlgtatat	ttgaaatttg	ccaagagagt	agatcttaag	tgttctcacc
25021	aaaaaattta	aaaagtaact	ataagagatg	actggttaag	taattagctt	ggcagtagtg
25081	aatatttcac	agtgtatat	aaaacatcat	gttgtaggtt	ttatacaatt	tttatttgcc
25141	aattgtacct	caataaagct	ggaaaacagt	atgtcagggt	attactctgc	tcaaaacctc
25201	gcaattgctt	cccttttcac	ttgtagtaaa	agctttatct	ttaggtttaa	gggccataca
25261	tgatctggcc	actactcttc	ccacctgccc	taactattct	gacacctcat	ctcctactta
25321	tccccttctt	tctcatataa	atccagtcac	attggcctgt	ctgctcttcc	ttgggccaca
25381	ctaggcatac	tcccattttca	gagccatgta	cttgctggcc	cttctgtgtg	ggatgctctt
25441	ccccagaaaa	tccacatggt	ttgcttccct	actttcttca	agtatttctg	ctaacagcac
25501	catcttcagaa	gtgagggcta	tttgaccact	ctcaccttct	ggcaattttta	tactctcttt
25561	ccctgtatta	tttttctctc	acagcaatta	ttacctttta	ctatactaca	tgattaattt
25621	ctgtctgtct	acactcacca	gaatgttgac	tggtatatcc	catcccccta	cccactgtct
25681	agcaaaaaaca	ctttgtgtct	aaaagagAAC	cttgcacaca	gcagatgctc	aatatgaact
25741	aaataaatga	aagaatgaat	aaatgtacca	ggtaagaatt	tagtaatgaa	aaaaaaataa
25801	ccaccctcac	aaaatttctt	gaatttatact	taggactcga	cttcccaact	taattgaatt
25861	aatacaccaa	tgtagaccac	taagtaaagt	cagagagaaa	atactggcag	atttaatgat
25921	gagaaatgca	ttggagccta	caagatagaa	aaactctgga	ttacttcacc	atatgaggga
25981	gctgaattat	tctgttagat	acctaatttc	tcttaagttt	tgagattaag	ataagagaaa
26041	atcaagccta	tcatttgagt	aataatgcaa	ttttgaaaag	tacactgctc	agaagtttaa
26101	aaatttctgg	tacagtccca	aactgaagac	ccactttttt	cttatctcct	gttgcttcag
26161	cttactaagg	cagagattac	tctccatttc	tacactatat	ctgatgtatc	taaattttgt
26221	ttatatgtac	ctactttata	tgtagtagga	aaaagatttt	ctaaaataag	ataYattttg
26281	ttcaccaagc	aaaattcaat	tcagtaatga	gaatgtcata	cctttttcca	catgagtttt
26341	gatcccagct	cttctctccc	tggtttctct	ggccatttct	ctaagtttct	cttctgtttt
26401	ttccttttct	ttctgagcca	tttttctctc	tacttgggca	cgcattttcca	cagcttcacg
26461	agcctgttgg	taaagtcaca	tgttaaacag	gacactatca	tgggataaat	atttatggct
26521	atcatccaac	tccctaacat	ctggctggta	agagcaatta	tacagaattc	agatttggtta
26581	ttgctaatta	aaacaaagag	tcttggtccac	taaaagata	cctgcctgga	agacactgca
26641	aggtcagctt	ttttctctgg	gaatattttta	gtgattgaaa	ctatcgtgac	cccaattaat
26701	cttgtaagtc	aagtattttca	ctaattgcttt	ctcagaaaca	atctaaagaa	caaaatttaa
26761	gccaggcata	gtggctctct	cctgcaatcc	tagcactttg	ggaggccagg	gtggaaggac
26821	tgcttgagct	caggagcttg	ggaccagcct	aggcaacata	gtgagatcca	tctcaaaaaa
26881	ataacatttt	caaaattaat	aacactttca	aagttaatca	agtcaactca	agataaaatc
26941	ttaaacttaca	cttgcttagat	aatttttata	agggtactca	ctgtgatata	aaagacattt
27001	tttttatggt	ccaaacacat	gctatgatgc	atgtgaactt	caatatcaca	tattaatcc
27061	ataatattta	tattttacata	tattttttga	gatgaagtgt	cttactctgc	tgcccaggct
27121	gaagtgcagt	ggtgcgttct	tggtcactgc	caacctccac	ctcctgggtt	cgagagattc
27181	tctgcctca	gcctcccaag	tagctaggat	tacaggtgca	tgccacatcc	agctaattgt
27241	tgtattttta	gtagagatgg	agttttcaca	cattgggtcag	gctggctctg	aactctcaac
27301	ctcaaggaa	ccaccagctc	tgccctccca	aaatgctggg	attatacaca	tgagccattg
27361	tgcccagcct	aattccataa	tattttaatac	attatgtaac	attaatagat	ggccaatgct
27421	atgaacgcca	ttttatgtaS	tattcaattc	attagaatgt	ttatctgcct	tttctgcaa
27481	ctgccctttt	gggggacatc	tacatctaaa	atagaggaaa	atttgaaatc	aaaaatttat

27541	aaacctagca	atcttcaaaa	aatgggatca	tcctaggaaa	ccattattct	gacatgttgt
27601	gagggagccg	gatcaccact	gtggggtgta	tgtacacatc	tacaaaagct	gggagcatgc
27661	agagagctgc	caactagcag	ctgccatttc	tgacacacgt	accatgggtg	agactctctt
27721	tgggtgctcg	gtactatagt	taaattatca	gtaccttagg	taaccaagaa	tggatttttaa
27781	aagtaacaag	agcattaaac	aaattataac	aaaccaacct	tccgatcagc	aatgtagagg
27841	gcttctgcca	atttggcgaa	attttcattt	atgtgtactg	tctgtagtcc	tcttccatca
27901	gcagccagac	gtttgtctaa	tggaaattgta	taaccctaga	gtgaaagcag	acagaaaacc
27961	cagtcacaga	agtgtctcaa	ttaatcgtct	gaagttaa	taattttatt	tttttcgagc
28021	tggagtcttg	ctctgtcacc	ctggctggag	ggcaggggtg	ccatctcggc	tcactgcaac
28081	ctctgcctcc	tgggttcaag	cgagtctcct	acctcagcct	cccagtagc	tgggaccata
28141	ggcgtgcgcc	acacactcgg	ctagtttttc	tatttttagt	agagacggag	tttcaccatg
28201	ttggccagga	tgggtttgaa	ttcctgacct	tgagtgatcc	gcccttggcc	tcccaaagtg
28261	ctgggattac	aggcatgagc	catcgagcct	ggctaaaatt	aaattttaat	atgccacat
28321	tttgagaacc	acttaactgg	aggttctttt	tccttgcatc	tttcttcaca	atgatgcaaa
28381	acaacagcag	ctatcaatga	gtgctttatg	tagaagatca	agcacttaca	cctatggtct
28441	catttaacct	tcacgataac	tctataaagc	ttctaacatt	attgtgtcca	tttcactgat
28501	gctttggtaa	catccccaag	gtcatacagt	tagtaagtg	gcgattagga	tttaaaactca
28561	Ncaggatttc	agaacccaag	ctctggaaaa	atagatattt	gtgttaataa	ttgtcgattt
28621	taataaaaag	aaatcactgt	caaacagttc	aacagttgct	cctcaagtaa	atttaacagt
28681	caaacaaaag	tagaatcact	ataatcttag	catctatagg	aagttaaatc	tgtgtactgt
28741	tcttaatgat	ctgttatatt	cttagaatac	tgaagtttta	gtaaaatagaa	aaaaaaggct
28801	gggtgcagta	gctcaagcct	gtaatcccaa	cactttggga	ggccaaggca	gggtgattgc
28861	ttgagcccag	gagttcaaga	ccagcctggc	aacatggtga	aaccccgtct	ctacaaaaaa
28921	tacaaaaaag	tagccaggca	tgggtgggca	cacttgtggt	cccaactact	cgggaggctg
28981	aggtgggagg	atcgccctgag	cctgggagg	tgaggctgca	gtgagccatg	atcataccac
29041	tgtactccag	cctgggcaac	agagcgagac	cgctc caaaa	aaaaaaaaaa	aaaaaaaaaa
29101	aaccaaagaa	aaaagaaaaa	aaaatgtttg	ctaaacaact	gaatttctac	atttctaaac
29161	tttgcacccc	atttcaataa	ctcaagaact	gtacttgaat	gtctatgtaa	ctttcaggct
29221	tagctttatt	ttccagatat	ctgactagtc	tctcaaatgt	tgacttttag	tccatggtct
29281	cataataaaa	tgaagcatta	tagaaccagc	tcaagatacc	atccagttac	aaaataccaa
29341	tcaactcaaa	tcagaccagt	cggcaacctc	tttgttttag	gtgagtatat	aaattctctt
29401	aagaaacctg	aaatggctct	ctttatccat	caaaaaataa	atcatgaact	atgattcagg
29461	gaagatgaaa	cagtaactag	caaactctatc	ccacaatcta	cagattctta	acaactccta
29521	gaaacttttg	caattattag	agatatatag	cttaagcctt	cactactaca	gagcccagta
29581	gttgaatatt	aaagcactta	aagtatccaa	gcatttactc	tgttcttttg	tcatataatc
29641	gtttacatca	gtttactttc	ttttacgaat	aatttttaga	tataaatttg	agttttctaa
29701	ataccatctt	gtattttat	gacactttat	tctacttatt	ccttctctat	gtgcttctgc
29761	tagctacaat	tcactaaaag	aactgatatg	gggcccgggtg	cgggtggctca	cacctgcaat
29821	ctaagccctt	tgggtggccg	aggccggcgg	atcacaagg	cgggagatca	agaccaacct
29881	ggccaacagg	gtgaaacccc	gtttctacta	aaatacgaaa	aattagccgg	acgtggtggc
29941	atgcgcctgt	agtcacagct	actcaggagg	ctgaggcagg	gcaatcgctt	caacctggga
30001	ggtggaggct	gcagtgaact	gagatcatgc	caccgtactc	cagcctagt	acagagcgag
30061	actccgtctc	Naaaaaaaaa	aaggaaactga	cacggtattc	acgtgcctgg	catactaaca
30121	tttctattac	attaaaaaag	tcaccaaaatt	gggcatctct	ttttgaaaa	gctcagttaa
30181	accagtggtt	gtttcaacta	cagttatgca	ggctgtagtt	gttttactgg	tgtttccag
30241	tggtttatca	taatagagac	tgatttgaca	acttaattac	ctttgcattt	ttccagttag
30301	aaatacaagg	aggaatcttc	cactcttgtt	gttcctttac	agtcatctga	gagaagaggg
30361	aaagaagaga	caagttttaga	tttatagctc	acaaagtaaa	tctaagtata	atctttatct
30421	tcaaagacag	ctccatatac	actgacatat	gcacagccgt	agtaaaaagg	cagcagttcca
30481	tttgtttgct	agcactagtt	taccatataa	ttctagtaat	aactttgaac	tagtgtagaa
30541	aatggaaaat	aatctagaaa	ttattcttga	cagttataaa	tgttggctag	tatttttcca
30601	tattttaaagt	gatgtgactt	ttttcttttt	tttattgaga	cggggtttcg	cttttgttgc
30661	ccaggatgga	gtgcaatggt	gcaatcttgg	ctcaccacaa	cctccgcctc	ctgggttcaa
30721	gtgattctcc	tgcctcagcc	tcccaagtag	ctgggattat	aggcatgcgc	caccatgccc
30781	agctaatttt	ttttaaagta	aagacgggg	ttctccatgt	tggtcaggcc	gggtctgaac
30841	tcccgaactc	aggtgatctg	ccgcctcgg	cctcccaaa	tgctgggatt	acaggcgtga
30901	gccaccacgc	ccggccgatg	tgacttcttt	tggatgaaaa	attgttcttt	tgattaacaa
30961	aacatatctt	aaaaagtcaa	attaccagg	acctaataaa	cccatatatt	cacttcttat
31021	aacatcctta	aactgaaaca	gcagaccctg	tccattccat	taaaagtagg	gtttgcatta
31081	ctaataacaa	aacttacagc	taacatgtta	tatacattct	cataagcatt	cttgaactt
31141	catcttgaaa	aataaaaaaa	gaaactacaa	gtcagtggtg	tgacatgcac	ctgtataggc
31201	atagggagg	tatctataaa	attaataacc	atgagtagaa	aaaaatattt	tggactacag
31261	caaagaatc	ctttcggcta	ggagaatgca	tgacaggcgc	aggaggagaa	gggtgtcccc
31321	ggggaatttt	cttattaatc	ctgaagtata	aaaacacaac	cagagagtga	tatagtttcc

31381	tccctttttag	tcatcatatc	aagtctctcc	tcccttttct	tgcataaaag	tttcaaacat
31441	agacaaaaac	agagactgta	taatgatata	catttaccac	tccactttga	caatgaataa
31501	aacatggcca	actttgaaac	tgttactcac	cccacctcca	caacattatt	gtcattttgc
31561	agagtaatat	acttttttcc	actttaaaaa	gcagctttag	gtcgggtgca	gtgggtcatg
31621	cctgtaatcc	cagcactttg	ggaggccaag	gtgggcagat	cgccctgacgt	caggagatcg
31681	agaccatcct	ggctaacatg	gtgaaaccct	gtctctacta	aaaatacaaa	aacttagctg
31741	ggcattgtgg	tacacgcctg	taaWccagc	tactcaggag	gctgagacaa	gagaatcgct
31801	tgaacccagg	aggcagaggt	tgcaagtggc	tgagatcatg	ccactgcact	ccagcctggg
31861	tgacagagcg	agactgtctc	tcaaaaaaca	aaaacaaaaa	caaaaaacaa	aacaaaaaaa
31921	ggcagcttta	ttgagataat	tgacatttat	tagctataca	tatttaaagc	atagcatttg
31981	attaaatttt	gacacaaata	tgccactgtg	aaatcatcat	cacaagcaag	atagtgaaca
32041	tgtctatcat	ccccaaaagt	ttctgcttgc	cccttgccca	ttctcttctg	ccctcccatc
32101	tccctctctg	ccctacctcc	ttcccagag	gcaaccacct	atctactttc	tgtaactata
32161	ggtttagttg	catttccctg	agtttttagat	aaatggaatc	atactgtatg	tacttgtttt
32221	tgccctggctt	cttttactca	gcataattgt	ttgagattca	tccatgttat	tggtgtgatt
32281	agttagctat	tccctttcac	ttacttgaac	cttgagggtc	ccattggatc	ttcttgcatt
32341	tctaccatcc	gaataaccct	ctgttttagct	ccagagttga	atgccactcc	ttgctgagat
32401	gggtgtgtatc	tgaagaaagc	agataaaaac	taaagggtgta	attagttcaa	gctttcagta
32461	ggtatttttcc	ttctattaca	atcttaagct	agactacatt	ttcaattcct	ttcgaaaagt
32521	ataaacagaa	aattaacaaa	atattttagac	aaagcagaca	gcaagaggga	taaaagaaca
32581	taaatttatga	gtttacagtca	aaatataaac	ttgagtggta	ttagtttgca	aataaccctc
32641	cctcatatgt	tttccaacat	ggaagggtgcc	gtgttcagtt	tcacagataa	gcagtctaga
32701	aaatagtacc	tatttagtat	gcctgcactc	gaggaaaaaa	cactcataat	aatcccagca
32761	ggcaacttac	ggtttagcagt	ttttgattaa	tcataggtta	taaaacagag	aaaagaataa
32821	tgacagcagc	cctaccagct	ctgaattttt	cctcaatatc	tgtatgtgat	tcaattctct
32881	atgcagattt	agtcagataa	aacgcctcaa	aatttagaag	tattataaag	atatctatta
32941	cttataaatg	gaaaactata	aataataatt	gtgatttagg	tcataataca	caagtaactg
33001	ttaatttYta	aaaacaaaaa	atataatttg	aaaatagcat	agggccaggc	cggttggttc
33061	atgcctgtaa	tcccagcaca	ttgggaggcc	gaggtgggca	gatcatctga	ggtcaggagt
33121	ttgagaccag	ccttgggcaa	cggtgtgaaa	ccctgtctct	actaaaaata	caaaaattag
33181	ctgagcatgg	gggtgcaagc	ctgtaatccc	agctactcag	gaggttgagg	cagagaatc
33241	gcttgaaccc	aggaggcaga	gggtgcagtg	agccaagatt	gtgccactgc	attccagcct
33301	gggtgacaga	gtgagactcc	atctcaaaaa	aaaaaaaaaa	aaaagaaaat	agcataggct
33361	gggcacagtg	gctcacgcct	ataatcccag	cactttggga	cctgaggagg	atggatcatt
33421	tgagttcagg	agttcgagac	cagcctggcc	aacatggtga	aacctgtctc	ctactaaaaa
33481	aaatacaaaa	attagccagg	tatggtgggt	ggcacctgta	atcccagcta	ctcaggaggc
33541	tgaggcagaa	gaatctcttg	ggaggcagag	gttacagtga	gccaagatca	caccactgca
33601	ctccagcctg	ggcaacagag	caagattctg	tccccacccc	caaaaaaaag	cataaaaaaga
33661	ggccaagtct	gggtggctcaa	gtctgtaatc	tcagcacttt	gggatgacaa	ggcaggagga
33721	ctgcttgagc	ccaggagttc	gaaaacagcc	tgagcatcac	agcgagaccc	tgtctctaca
33781	aaaaattaaa	atattagcca	ggcatgggtg	catcaaggca	ggaggatcac	ctgagcccag
33841	gaggtggggg	ctgcagtgag	ctgtgttggt	gtcaccgtag	tccagcctgg	gcaacacaga
33901	aactttatct	cagaaaacaa	acaaaaacca	aagtggacac	cttaaaaaat	tctgctttca
33961	attgtctgca	aagacagtag	acaggctttg	agtcattatt	gaaaacaatg	catttataat
34021	ggttgctaag	tgggttagtg	aaagaatacc	atgaccccc	aagctgaatc	aagttcagaa
34081	tgaaactaag	tcttctaggc	atgattcata	ccggatatac	tgagcaggag	ccaatttgct
34141	agctgctcga	actggcatgg	ctgcggcgac	cttctgtgat	acagattttt	ctaaggctac
34201	tcttgtcttt	tctgttatct	gaaataggaa	atatcacatt	gagagtttgg	aagttaatca
34261	ggatgtaaat	agtcacagaa	tatatcaatt	cccagttacc	tctttaatag	cttcttcatc
34321	gggccttttg	aggtctggat	catctgcatt	cataacctcc	tttggaacca	ggtcagtgtg
34381	tttgctataa	atgacctaaa	atggttcaaac	acaagcaggc	tttaagaaaag	caaacacaga
34441	aaagaaacct	cattagccat	ttcagatgtc	aacatgccct	ctcagcacac	ttaaaaattt
34501	aggacactca	gattttctct	tttagatttt	aaaaccaagg	aagcttaaaa	tatatataca
34561	aattgctacc	ctggctgtaa	gttagtaatt	ttcccttgac	ctcatgcatt	taaattgttt
34621	gagactttct	aatcaagtta	gtatccaaat	aaaccatgcc	aacaaaaatc	tgcatgtatg
34681	atcactattt	ttgaggccca	ctaaaacttg	aaattaatga	taatacaaat	tgccctatag
34741	accattaacc	tggattacaa	actcatcaga	atttaagtca	ggttaccaaa	taattttcaa
34801	attcactggg	cctgaaaaat	tttattgacc	tcactcctatt	tataacgctt	ctcttttctg
34861	ccatcaactc	tattcacaga	ctaagaaaaa	aaacatgtgc	cacagttaaa	aaaccagaaa
34921	tgcctttctt	catgagaaga	caaaaagaaa	aaaaaaagaa	aaagaaaaat	atatcttata
34981	tatacaactt	attttatatc	cattatccaa	attgaggcct	gttatatgac	ttctactatg
35041	cccaacatgg	tttctattag	gctgacactg	atgtggtagg	ctcttaacaa	atagaatact
35101	ggttactttc	aggtccatct	atacttctta	aatattttca	taacaattta	catggcatat
35161	tttaaagtaa	aaaagagata	cagagggggc	gggcgcgggtg	gctcacgcct	gtaatcccag


```

35221 cactttggga ggctgaggtg gctggatcat gtggtcagga gatcgaaacc atactagcta
35281 acatggtgaa acccgtctc tactaaaaaW acaaaaaatt agccagttgt ggtggcaggc
35341 gcctgtagtc ccagctactc gggaggtcga ggcaggagaa tggctttaac cggggaggaa
35401 gagcttgagc taagccaaga ttgcgtcact gcactccagc ctgggtgaca gagcaagacM
35461 ctgtatatatta aaaaaaaaaa aaaaaaaaaa aagagagaga tacagaaaat tcaataaata
35521 ataggcattc tacctttaat gcttttagtca gaaattcttt gaggtacata aaaagaaagg
35581 acaacttata aatatcaaaag aaaaatcaaat tcctaggcac ttgagatctc cagaattttt
35641 ttttgtagaa ctgtttacca ctacaagcca aagtcttata ttaataagtc tatcatacat
35701 cattcattaa aaaagtttaa aggaaatgca tgctatttct aattttctac tocaaacaga
35761 gcaaaaaagt gtgagccact ttatggctta tgcaacttga cctgactttg cctatgcctt
35821 caatctcccc ttacaactct cactctctca ctctattatgc ttcagtcaca ctgactcatc
35881 ttctgtttct tgaacataac atacttttcc ccaattctag gcctttacag agttctgttt
35941 atcattcttc aagtaagctc tcaatctgtg tgtctcagat caaaggctct ccattatcga
36001 ggcttcttct gattgcctta gtagtaaccc ctccccaccc aatcatttgc tocatagcac
36061 tgtttactgt gatcaagact gttcacagca tgaaattatc tactttacat ggttattatc
36121 tgtctcccgat attaaaacaa taaaaacggg gagggcacgg tggctcacgc ctgtaatccc
36181 agcactttgg gagggcggagg tgggtggatc acctgaggtc aggatttcaa gaccagcctg
36241 accaacaagg tgaaaccccc tctctcccaa aaacagaaaa attaacagg catggtgggtg
36301 ggtgcctgta atctcagcta catgagacgc tgaggcagga gaatcgcttg aaccaggag
36361 acggagggtga cagtgaagcg agatcatgcc attgcaactc agcctgagca acaagagtga
36421 aactgtctcc aaaaaaaaaa aaaaaaaaaa aagtgtctcc tgtgctttat aataatgact
36481 ggagatattt gatcaatatt ttttgtacaa atgaataaat aaatgaattt tagagatatt
36541 ctgttacatg ctaaagggtgt ttataggcca tgaacagtga ctacgcctg taatcacagc
36601 attttggaag gccaaagggtg gaggatagct tgagccagg aattcgagac cagccagggc
36661 aacatggcaa gacctatctc tattttaaag aatggatgga tgaatgaata aataaataag
36721 tttatatattt aggcggtaca gaactttcta taagctactg atttccaga aaaaactatc
36781 tcatggctat acactaaaaat ctccagcttg ttttctcag gagtgagaag agggatcagg
36841 taagaaagcc taagatgatt tagcacaaaa cacgtactac tgatatgtga gttattagtg
36901 tcaccatact ctaagagtgt tgattttgga agggaccata aagctcactc tactgaaaag
36961 agtgagggttc agtacattaa gatcacacag ctactatctg acctcaaaag tcttatttcc
37021 tctcagcatg gaaaaaccta tacaacccta agctgaaaaa tatataacaa cacaagaaaa
37081 aaataccaat tctatgaaga gttaatcatg aRgagtttat cacttctcca tagcaacggg
37141 tataaggaaa ggattcctaag taaatggctc cctggtagtc atgtatgttc aacataaagt
37201 gtcaataacta attttagcata caagagtcat ttgctgctcc ttatttgggc attccaaagt
37261 taataaccct tcaaatgagc atcaRtaaga ttctagtatc tcacaaggct gagtgaagga
37321 catctgaaga gcaagattcc tccatagtaa ttaattataa aaccccccat tttttttttt
37381 ttgagacgaa gtcttgctct tgttgcccaa gctggagtgc aatggcgtga tctcagctca
37441 tagcaacctc cgcctcccgg gttaaaacca ttctctgcc tcagcctccc tcccaggtg
37501 ctgggattac aggcgcagta caccacgtcc agctaatttt tgtattttta ttagagatgg
37561 ggtttcacca tgttgccag gctggtctcg atctcctgac ctgagcgat ccaccgcct
37621 cggcctccca aagtgtggg attacaggcg ttacagcgag cacctggcca aaaccccaaa
37681 tttttatagt ctgagaatta aatgtgctga ggaaactgag taccaggagg caaatgggct
37741 ggggagYgtt ttctgaaaaa tacaatatac tcttgaaagt atcagaagca caaataagct
37801 ttgtcaacaa gataaagtag taactcatct gtgaaacaga tattttatta cagaaaactg
37861 gcaaagaact agactgacta aaaggataag cctgtcaaaa cattcactat gtgaggtaga
37921 aagaaaacgc taatctagct tagctagtgt gtgaaatcat acaaacaaaga taacatataa
37981 tttatcatgg ggctcagttta aagttgttac ggttttttag gggtaatacg ttgaaggcat
38041 tcttttgggt atagtaYgtg tgtaacattt cgtttcttta agagataggg tttcattata
38101 ttgcccaggc tagtctcaaa ctctgtgccc caaatgatcc tcccaccta accttcccaa
38161 agtgctggga ttacaggcat gagccaatgc acttggcctg tatgtagtat ttcttacata
38221 actatttcag agggaaaaat acatacacca agccctttat ttcaagattt atcacttggc
38281 tgggtgtggt ggctcatgcc tghtaatcca gcactttagg aggcgagggt ggggtgatca
38341 cctgaggttg ggagttggag actagcctga ccaacacgga gaaacactgt ctctacaaa
38401 aatacaaaat tagtcaggcg tgggtggcaca tgctgtaat cccagctact caggaggctg
38461 aggcaggaga atcacttgaa cccgggaggc agaggttgcg gtgagctgag atcgtgccat
38521 tgcacccag cctgggcaac aagagcaaaa ctccgtttca aaaaaaaaaa aagatttatc
38581 actttcataa ttaggaaaaa aaaagcattt ggttgatatca agatttacac taaaaaaaac
38641 tttttcatcc ccaattttgt atgactaata agtcgtatca acagtactta cttcacatgt
38701 agagaattgc cttagacttg tcctagaata ttacgtctta gaatcctcac tgattcttct
38761 caccctttat acctaacacc tagtattttc cattcttact cattcactat taatatttca
38821 gtgagaactc cctcgtttct gaagaatacc agacctctMR gaagcaaatc agaaattctta
38881 gtccaaattc ataaatccaa taacctgttc agtctctgga ataacaaatt ttttaaatcc
38941 tatttttatat attatgcagt ccatctctat ttcttctaaa acaggaaaat cagtgtgtgt
39001 caaacagggc atcactgcca cctctaactc ctctctccca gtgcctttgt aaatgtgtga

```



```

39061 ggtgttttcta gttaagacaa tgggggtagg atgggggacgc tactgggatt tataggtagt
39121 gccacttatg ctaaacattg tacaatgctg aagatggtcc taaacaatca aagtgtccc
39181 ccagtgccaa tggcacataa tagttacttg attaatgttg aattactcaa tgcattccaa
39241 acttcatgca tgaatctatc cacttctaga aaacaaacaa acaaaaagat ataaacctga
39301 gagtgggtccc agtttgggggt ttgtttctag tggctgcatt acagacaaaa gcaacgatct
39361 ggccgggtgt agtcgctcat gcctgtaatc ccatcacttt gggaggttgc agtgagccaa
39421 gatcggtgcca ctgcatttca gcctgggtga cagagtgaga ctccatctca aaaaaaaaaa
39481 aaaagaaaaa aagaaaatag cataggctgg gcacagcgac tcacRcctat aatcccagca
39541 ctttgggaag ctgaggagga tggatcattt gaggtcagga gtttaagacc agcctggcca
39601 acatggtgaa accctgtttc tactaaaaat acaaaaaaat tagctgggca tgggtggtgtg
39661 cgctgtaat ccagctacg tgggtggtg aggcagaaga atcacttaaa cctgggaggc
39721 agaggttgca gtgagccaag attgcactcc agcctgggca acagagcaag tctccatctc
39781 aaaaaaaaaa aaaaaaaaaa gaaaagtaaa gaaaagaaaa gaaaaaaaag agaaaaaca
39841 aggaacaatg ctttgggagac cttatttggc caagatacta aactccaaaa atagatagtt
39901 cagtctgaac agttctactc tattttgaat acaaaataat caccaatata gacagtgatt
39961 taaacatcat tccctagaaa tcagataaaa ctctgtcttt ccagtcttga aatataatgc
40021 aaaacggacc aaagctgtat ttcatcaaaa ataacacaac ctgccaatgc caaattgttg
40081 cccttagatg ttgttttttc ctaccaaag ggactggaga ggtagcacat accaataaag
40141 aagagtgatt acccccaagg agaaatgcac caaaacaaag gcaactgaag tcaagtgaag
40201 gcctccatgg tatagRgtaa atttccagta tatgttctcc atgtataggg gcctaatttt
40261 ccttttgagc gagtggacta Yacttctcca ctggcttcag aagtggtttt gtgttataca
40321 atggaaacca ctattatttg ctgaacaaac actggtggac gccttatgtg tcaaactttc
40381 aggcagtaag taaggcacca aagacggaaa tataaatata taatccctgc ccttaaaaaac
40441 attatagtca aatgatacag attatgagtg aacagataat tacaacaaaa ggtgaagact
40501 gctatgataa acattgagag tagagaagag ggacactcag tctagccagg gaaggtccag
40561 gcaaacacc tggaatagtg aagagtggag tgagaatgga aagcaataga attaagtcag
40621 agaatgaata agggtagaac tttctaggaa aaagaatata catgggaagt ctcaagagag
40681 aatgtgtcat attatgggaa ctacaagtag ttcctcatgg caaatattta tagggaatgc
40741 ttgaatgttt tttcttcttg gctgggcacg gtgtctcatg cctgtaatcc cagcactttg
40801 ggaggccaag gcagggtggac tgcttgaggt caggagtcca agaccagcct aaccaacaag
40861 ttgaaaccca tctctactaa aaagacaaaa acgagctggg catggtggcg gcgcctgta
40921 atcccagcta ctcagaaggc tgaggcagga gaactgcttg aactgggaag gctgaggtcg
40981 cagtgagctg agattgccct actgcactcc agcctggggg acaaagcaag atttggtctc
41041 agaaaaaaag tttttttctt cttaaacaacc accttggtca ggtgcagtgg cacacaactg
41101 taattccaac actctgggag gccaaagcag gcagatcact tgagctcagg agtttgagac
41161 cagcttgggc aacaagggtga aacctgtgca ctacaaaaaa tacaaaaatt agcgggcat
41221 ggtggtgtgc acctgtagtc ccaactcctt agaaagctgg ggtgaaagga tcaacttgagc
41281 tcaggagttc aaggctgcag cgagtctgta ttgtgccact gcaactccagc ctgggtgaca
41341 gagcaagact gtctcgaaaa caaacaacaa cacatataat agtatattat ttatagttaa
41401 acaaaacaga atggctaaaa ctgaactatt ttttttggg tttcagatatt ccSatcagga
41461 gagttgaagc tagccaagaa aagttagtaa ggggtcaagc atgccatagc agcttagtcc
41521 ttaccctaaa gaccatggaa gccatagaaa gattttaatt gtgcacactt gcgtgtatgt
41581 gtatatacat gatgagaaag acaaacagac atggatgaat ttgtttgttc atgccagcaa
41641 gacaaatgtc aatgagatgg ttgcaggctt ggctgcagta aatcaagcaa gaaatgaaaa
41701 cagcctgaac taagtagcag agaggggaat atacttgaga agtgtaaagg aaataagatc
41761 gttcacagag aataaataga tggaaaggtc gagggaaaaca agaactagaa taattgctga
41821 gcagagagaa catatgggca gtatgtgtca catgtgaaat ctaaatagcc tttgagccag
41881 taagtccaca cctaagaatt taaacaaata atcatgttca acaagatgtt tgtatgtgga
41941 tcttgaatac atttgtttgt aaaactaaaa accaccaatg cttaaaaaaa ctagcagata
42001 accatgccac ggaatcctat atagattaat aaaaaaaaat gaagtagact tatatgtact
42061 aacaaaaatt gtcaaatgtg ttaagtaaaa Waaaaagcaa gtcaaaggaa gttatgggca
42121 tcaaaatttc attcatgtaa acaacagttt ctaagtttgt atgtatacat aaaatgagga
42181 tatacataaa tatctaaatc cataataaaa atttggaatt atagctccc ctccccctcc
42241 ccccttctgt gcacgggtctc cctctccctc tccctctcca cggctctcct ctgatgtgta
42301 gccgaggctg gactgtgctg ccacctctc ggctcactgc aacctccctg ctgattctc
42361 ctgcctcagc ctgccgagtg cctgggattg caggtgcgcg ccgccacgcc tgactggtt
42421 tcgtattttt ttgggtggaga tggggttttg ctgtgttggc cgggctggtc tccagctcct
42481 aaccgtgagt gatctgccag cctcggcctc ccRaggtgcc gggattgcag acggagtctc
42541 actcactcag tgctcaatgt tgcccaggct ggagtgcagt ggcgtgatct cggctcgcta
42601 caacctccac ctcccagccg cctgccttgg cctcccaaag tgccgagact gcagcctctg
42661 cccggccgcc accctgtctg ggaagtggag agcgtctctg ccaggccgcc catcgtctgg
42721 gatgtgagga gccccctgc cccggccgcc agtctgggaa gtgaggagca cctcttccc
42781 gccgccatcc cgtctagga gtgaggagcg tctctgccc gccgccatc gtctgagatg
42841 tggtagcgc ctctgccccg ccgcccggtc tgggatgtga ggagcacctc tgcccagccg

```

42901	ggaccctgtc	tgggaggtga	ggagcgtctc	tgccYggcca	ccccgtctga	gaaggggaaga
42961	gcccctccgc	ccggcagccg	tcccRtctgg	gaagttagga	gcRactccgc	ccggcagccg
43021	ccccgtccag	gagggaggtg	cggggcagcc	ccccccccgg	cagccgcccc	gtctgggagg
43081	gaggtggggg	ccagcccctg	cctggccagc	cgccctgtcc	gggaggtggg	gggcgccctc
43141	gcctggccgc	cgccccatcc	gggaggtggg	gggcgcctct	gcccggccgc	cccgtctggg
43201	aggtgaggag	cccctctgcc	cagaggccac	cccgtctggg	aggtgtaccc	aacagctcat
43261	tgagaacagg	ccatgatgac	Ratggcggtt	ttgtcgaatg	gagagggggg	aaatgtgggg
43321	aaaagataga	taaatcagat	tgttgctgtg	tctgtgtaga	gggaagtaga	cgtaggagac
43381	tccattttgt	tctgtactaa	gaaaaattct	tctgcctttg	gatgctgtta	atcataacct
43441	tacccttaac	cccatgtctc	ctgaaacatg	tgctgtgtcc	actcaggggt	aaatggatta
43501	agggcggtgc	aagatgtgct	ttgttaaaca	gatgcttgaa	ggcagcatgc	ttgttaagag
43561	tcataccacc	tcctaatact	caagtaccac	gggacacaaa	cactgcggaa	ggcgcagggg
43621	tcctctgcct	aggaaaacca	gagacccttg	ttcacttggt	tatctgtctga	ccttccctcc
43681	actattgtcc	tatgaccctg	ccaaatcccc	ctctcYgaga	aacacccaag	aatgatcaat
43741	aaatactaaa	ataaataaat	aaataaataa	aataattttag	gaaaaaaaaa	tttggaaatta
43801	tatatatcac	tgtagtaaca	gcagtacact	ctgggaagta	aagtaggaga	aaagcgggca
43861	ggactRtata	tacttgtgat	atatacttat	ggtatctttt	actataatat	attattttga
43921	taactttttt	ttttatttga	gacagagtct	tgctctgtcg	cccaggctgg	agtgcgggtg
43981	tgccgtcttt	agctcactgc	aagctctgcc	tcccaggttc	acgccattct	cctgcctcag
44041	cctcccaggt	agctgggact	acaggtaccc	gccaccatgc	ccagctaatt	ttttgtgttt
44101	ttagttaga	cggggtttca	ctgtgttagc	caggatgttc	tcaatctcct	gacctcgtga
44161	tccaccacc	tcagtctccc	aaagtgtctg	gattacaggt	gtgagccacc	acacccagcc
44221	tatttgtata	attKttttta	aaaaaatcaa	tataatagta	agaaaaaatg	actcccaggc
44281	tttcagcata	aacaattagg	aaaatatcac	taccacttac	tgagttacag	aaatagcagg
44341	tagattaagt	ggaaaattat	gtattttagt	ttggatttga	ttaagtagca	ctttagggaa
44401	atctagatac	acttgccctt	gtagacagta	agatataagg	acctacagcc	cagcagagaa
44461	tactgggctg	gaaacacaga	ttgaacagtc	atcagcatat	gagtgccaaa	taaattacat
44521	tagttaaaga	gaatgtgctg	tgtttttttg	ttaaagtcac	agaatgatta	aaagaaaaaa
44581	atatataaag	agaatRtgct	gtaaaaaaat	gtgttaaagag	acttttgcca	gagagcacct
44641	catcaaactg	gttttaggtga	ggtcattttg	tagtgaattg	aagagtgaag	ggaaagtga
44701	gaacaggggc	aaaaacaact	ctgtaaacaa	gcctgaaaaa	agtgaagaaa	cttaggtggt
44761	agctagacag	aacactgcga	gtaggtactc	agggaaaagc	ctatctttca	aaagataaga
44821	aaggccaggc	agcctcatgc	ctgtaatccc	agcacttttg	gaggccgagg	Ygggcagatc
44881	acgaggtcag	gagatcaaga	ccatcctggc	taacatgggtg	aaaccccatc	tctactaaaa
44941	atacaaaaaa	attcacccgg	catggtggca	ggcgccctgt	gtcccagcta	ctcggggggc
45001	tgacRtacga	tgatcgcttg	agcccaagag	ttcaaggttg	caatgaatta	tgagggtgcc
45061	agggcactac	aatctaaaaa	agaataaaca	aacaaatgac	aagaaaaaca	agcacattca
45121	aatgctaaca	ggaaagagcc	gttagaagag	gagggatgaa	gaactgttga	gaagattaac
45181	tagtaggtaa	ggttcataga	ggaagagggg	agggattggg	atccagaata	catatggcaa
45241	gaatcaggtg	agatccaggt	gcagtggctc	acgcctgtga	tMccagcact	ttgggtggca
45301	gaagcaggag	gattgcttga	gcccataaag	tcaagaccag	cttgagtaac	atagggtat
45361	cccacctcta	caaaaaaata	ataaatttag	caggcatggt	ggtgcacacc	tgtgctctca
45421	gccactcgag	aggctgaggt	gggaggaatg	cttgagccca	ggaggtcgag	gcttcagtga
45481	gccatcatcg	tgccactgca	ctccagcctg	ggcaacagaa	caagaccctg	tctcaaaaaa
45541	acaaataaaa	aataaaataa	aaaaagatta	ggcttagaca	gaagaaggca	tttcgtgcaa
45601	aaacaaaagt	aaaagtcgag	gaaaaaggac	aagtgcagga	tgaagaaaat	tgacaattgt
45661	cttccaatag	tctttatatt	caaatagatg	ggggagtcac	ctgctaagga	tgagatggga
45721	tgaaaggagt	aggaattaaa	gtttggtggg	agtaccaagg	gagctgagct	gaaataggag
45781	atataaatct	gcatgattgc	ataattgtct	tccccaaaag	ctcagcagcc	atgatgcaga
45841	agcaaaggat	gggttggttag	tgctatttta	gagggtaaat	tgaaaaagtt	tggcggggag
45901	agagaggcca	atgataggca	ctgaaagctc	tcaagtattt	ttccaaacat	catcagttct
45961	aaggaaatct	ttcacttcta	tcttgagtat	ttatttaca	gaaaaaaaaa	aaactaggtc
46021	ttcttaaaaa	tagttccctg	gccctgagtg	tagctgttct	gtgatgagat	tatttttaagc
46081	aaattttatt	tttggcctat	ggattacttc	gtagcagggc	catcctaata	ctggatagca
46141	aatgctagtt	ttgtacttaa	tctagggcaa	aagaagtttg	tgtaatttat	tatactcaaa
46201	taaaactatta	aatacttgta	atagagtaca	tgtaagtact	ctaagtacat	ggaggaatg
46261	agtcaagact	ataggaaaaa	actagaagaa	actacaggat	aaattaaaaa	ttgatggtaa
46321	ggctctgggg	ttgcagattt	ttaaaaat	actggcagtt	ttagaatgct	agtaataata
46381	agtaactggt	ttaaaatgta	cttttttccc	cctttaaaac	aaggtattgc	cacccaagca
46441	agagtgcagt	ggctcactgt	atcctcaaac	tcttgggctc	agtcaaacct	cctacctctg
46501	ccgtcccaagt	acgtgggaat	ataggtagca	gccaccatgc	ccagccactg	ctacttttaa
46561	cataattcaa	acagcaagag	atttatcggt	tccacaaaat	ttaaaatggt	tcctaaaaata
46621	tttatcattc	agggaaaaca	tgagaggatt	accttgtctt	ttgactgtcc	ttgtcgagca
46681	attgcatcat	atttaatttt	tccttcagaa	tccacctgaa	tggccagcgc	attcgacatt

46741	tttttctttc	gtcccatatc	cagtggatag	tggggccacat	ggatctctgg	aaaagcacct
46801	ccatctccaa	aatcctacac	attagaagta	gttaaataaa	atagttaatc	atcatttgac
46861	attcaagata	tattcattca	ttcaacaagt	aattgttgag	tccttagtat	gctagatgct
46921	agagatacag	cagagttccc	tgcctcttgt	ggagtttaca	ttctagtagg	caaagatata
46981	aaaaacacat	tctttggcct	gtgataagta	ccatgaaggt	aatataaaaag	aaaaagactg
47041	ggaaaagggg	ctacttttga	taggggtggt	ggggacagcc	tttgtgaaga	agtaacatgt
47101	gagctaaagc	ctgaatgatt	agaaagcagt	catgggaaga	cacacacaca	cacacacaca
47161	cacacacaca	cacacacMcC	acacacacac	acacacaaag	catttttctag	atggatcaaa
47221	tagccaatgc	aaaggccctg	ctgcaagaat	gagcttggtg	tatttttagga	aatgagaatt
47281	cagtggggct	aacagcttaa	taaggacagc	agagcttgag	gcacaagtta	catgagtaaa
47341	aaacaccaga	tcacaaaggg	acttacatgc	tacagcaaga	aaatgtggat	ttcatatatg
47401	caggacgaaa	ctaagtaaaa	tggctaggac	aaggcaaaat	atatgtgtat	atatttttagc
47461	atataaatac	atatttttagt	gaggaacagg	taaagctata	cttggggcag	aatataaaaa
47521	gatttttttct	tggaaatcaga	ttttattaac	atattatttta	ttcattgttta	tgtgagaagc
47581	aaatgaatac	aataggaagg	atccagaaat	gagtttacta	cattttctact	agtgattata
47641	cataatttatt	taggagagag	aactgagact	gctaaagtag	aacaagctaa	ataaagtttg
47701	ctgatgaaat	gtagtatttt	aaaaaggttt	atcaataaaa	catgaggatc	aagtataatt
47761	tttaagtagt	taaaaaaaat	tcagttagtat	ttgcttagta	cacttttaaac	atgatKcaca
47821	catgtatttat	ctcattttaat	ccttgtaact	ctttaacatc	attttaaatg	gaaacattca
47881	ggcttaRaaa	aattaagcgg	tgttcctgag	atctcatagg	tattatatgc	agcacctgaa
47941	cccagatctt	ttgagactcc	gaaaagcttt	gtattttcta	ctactatgaa	atacctttct
48001	tttagctcaa	aacagccaaa	tgagtactat	taagaaataa	tagaaaaacac	aaattaaaaat
48061	taatcatttta	gaggggggtt	gatggaaaaa	aaagaaaata	ttcactttta	caagacctct
48121	ttaagtaata	ctcYRttaga	tctaagatat	atgagttcct	aagtacaata	tgcttatttg
48181	ctaaagagat	gtttgaagtt	tttgaaaaacc	aagagaataa	agctgagaga	cattttcttg
48241	taaatgagaa	agagttaaag	cattcctaag	ataagKgctt	cccgtacca	agccatcaat
48301	ctgacacatg	cttggatttg	ctagagtttt	tcaactataa	taagtacttc	agaactgtag
48361	ctgtagcatt	atgtgggtcat	taagcttgcc	agaaaaatcaY	agcttttatt	ttgaaaacta
48421	gaacaaaatg	ccttcttcca	cagaacattc	actccctca	agcagcatcc	atagcatttc
48481	cgatttagct	agccattttca	acttgatttt	caaataagtt	gttctgtttc	accagatat
48541	atgaaaccac	tgatagccaa	gtaagaactc	aaaaatctaa	tagcaaccac	ttactcgtaat
48601	gctgaaatgc	atggatgaga	tttacattct	atgtaagatg	atgcctgtca	ccctcattgt
48661	gcaatgtact	gacagaccaa	acaatattta	taatgcctca	acagagacac	tcaactgctc
48721	tattagatga	tgccaatcac	tacttttttg	ccccaggtaa	gactccttga	taatctttca
48781	tcttctctctg	cctcccaYaa	atWacactat	cagagagaac	aaaagcaaca	aaaatccctt
48841	aataattgca	atatgactgg	ccttcacaca	ccagatttca	taaggaccag	aggaaagctg
48901	acaaaaaagc	aacaaaaaga	cctacatgtg	ggtttttttg	aatacatagg	acaagatatg
48961	atctagtttg	gtgttcctct	tataaactct	aataaagagt	gtggggcagg	cgcggtggct
49021	cacacctgta	atcccagcac	tttaggaggc	cgaggcagg	ggatcacctg	aagtcaggag
49081	ttcaagacca	gcctagcaaa	catggtgaaa	ccctgtctgt	actaaaaata	caaaaaatta
49141	gctggRtgtg	gtggcacgtg	cctgtaatcc	tggctacttg	ggaggctgag	gcaggagaa
49201	cgcttgaacc	caggaggcga	aggctgcagt	gagccaagat	cgcaccactg	cactccagcc
49261	tgggaaacgg	agcgagattc	catcccaaaa	ataaaaaaaa	taaaaattaa	tttaattaatg
49321	taaaggatcc	tgagtaatta	gaataggtgc	tgccaaggaa	gcttgttctt	tttattttta
49381	tttttttattt	tttttgagat	ggagtttttg	tgttgttgcc	caggctggag	tgcaattggtg
49441	cgatcttggc	tcaccacaac	cctgcctccc	caggttcaag	caattctccc	gcctcagcct
49501	ccggagtagc	tgggattaca	ggcatgcgcc	accacacctg	gctaattttg	tatttttagt
49561	agagacaggg	ttgctccatg	ttggtcaggc	tggtctcgaa	ctcctggcct	caggtgatct
49621	gcccgccttg	gcctcccaaa	gtgctgggat	tacaggtgtg	agccactgcg	cccagccgga
49681	agcttgttat	ttggcaacaa	gctagagata	gaaagagcta	ttccagatgg	ttccttctcc
49741	ctcattgcct	tcaaatcaac	caaattttgt	cttataccg	caaaaaata	tttaactctat
49801	atatcagata	tatattaaca	ttctcaaatt	aaagtaaga	aaacaggtga	agagaaagaa
49861	atcagcaaa	gaagtgttac	tggagttttc	cctgaaggaa	agttgtctct	tcagggttat
49921	gtggtattat	cggaggggag	gaggaggagg	aggagcccaa	atatgaaaWc	cactttcatc
49981	tctcttatct	tttactttcc	cattattttc	cgtaattcct	gaatttataa	aatattagtt
50041	ttcttgtaaa	tatttaatag	tttgatatag	gaatagacaa	atgggtacagt	tctctagtag
50101	ccacaaagca	gtgatagata	ataagcaaaa	agaaaatgat	acatcttctg	agggaaatat
50161	acatcttgag	agcacataac	cagaattttc	atcacactga	cattcataat	cagaggttct
50221	aacattttaa	tttatgttat	agtgtgctat	aataaatttc	agcacaattt	aacaaatcta
50281	aacttaagat	ctatatgtac	ctctaataac	cgaggtatcc	agcctttccg	gtatccgtac
50341	gggggaggtt	ctcttcggga	ggagaccagt	gaggtctgcc	gtgatctctg	ggatcttgcc
50401	ttttcttcag	cctcaagctg	gtcctgagat	agctgagtag	gtgcaggtaa	aaagctataa
50461	gcaaagataa	taaaagtcag	aaatttaaaa	aactcttatg	tttaatttga	ccactgcaac
50521	ttggccacag	agacaatttt	tcctacgatg	cttttacttt	tcagaaaaag	agcagatttt

50581	tcaatgattt	atggttagtac	tcaatactag	ccattttata	atatgtggaa	gacagaagta
50641	ctcaaagata	cacagtaaaa	acatctgact	aggacaacag	caaatagatgt	ccagaataaaa
50701	aaattaacag	atatttcaca	taacagacaa	actcttaagc	gtttatgctt	ccatattttcc
50761	tttttttttt	gagaaagagt	ctcactctgt	tgcccaggct	ggagtacagt	ggtgtcatct
50821	ccactcactg	caacctctgc	ctcctggggt	caagtgattt	tcctccctca	gcctcccaag
50881	tagctgggat	tacaagcatg	cgccaccaca	ccaggctaatt	ttttgcaatt	ttagtagaga
50941	tgggggtttca	ccatgttggc	caggctgggc	ttgaaactga	cctcaagtga	tccgctgacc
51001	ttggcctccc	aaagtgtctg	aattacagac	gtgtgccacc	acacctggcc	atgcttccRt
51061	attYccatat	ttctttcctt	tttttttttt	gagacagagt	ctctctgtgt	cacctaggct
51121	agagtgcagt	ggcaccatct	cagctcactg	cagcctccac	ctcctggggt	caagtgattt
51181	tcctgcctca	cctcccaag	tagctgggat	tacagatgtg	cagcaccacg	cccacctaat
51241	ttttgtattt	ttagtagaga	cagggtttcg	ccatgttggc	caggctgggc	tcgaactccc
51301	gaccgcaggt	gatgtgcccg	ccccggcctc	ccaaagtgt	aggattacag	gtgtgaacca
51361	ccacgcccag	cctgtttcca	tatttcttat	aacaaacaag	taaacattcc	taatacttca
51421	gggataaaact	gtccagtata	tacttgctat	tcttatactc	cttttttttt	ttgagacaga
51481	gtctcgttct	gtgcccagc	cttgagtga	gtggtgcaac	ctcggtcac	tgaacctct
51541	gtcgcttggg	ttcgtgtgat	tctcatgcct	cagcctcccg	agtagctggg	attacaggtg
51601	catgccacca	catccctatt	ttttgtattt	ttagtagaga	ctgggttttg	ccatgttggc
51661	ccatgttggt	caggctgggc	tcgaactcct	gacctcaagt	gatccatcag	cctccatctc
51721	ccaaagtact	gggattactg	gcattgagcca	ccacgcctgg	ctactccttg	tttttaacat
51781	ttttaccatt	atgtccctta	ctttaaacta	cttcagtgtg	aatacaaaaag	aactagattt
51841	aaaactgaat	gtaaaaaaga	ggtggataaa	ataaacacaa	gttgcatatt	taagtgtcaa
51901	gggtatctac	aatgaaattt	aaaacttaga	ggccaggcca	gccatggtgg	ctcacacctg
51961	taatcccagc	actttgggag	gccgagggtg	gtggatcact	tgaggtcagg	agatcaagac
52021	cggcctgatc	aacatggtga	aaccctgtct	ctactaaaaa	tacaaaatta	gccaggcgtg
52081	ctggtgtgcg	cctgtaatcc	cactcgggag	gctgaggcag	gagaatcact	tgaacctcgt
52141	agggtggagg	tacagtgagc	caagattgca	ccattgcact	ccagcctagg	caacaacagc
52201	aaaactccat	ctcaaacaaa	caaacaacaa	caaaaaaac	cttagaggcc	aaagaatcaa
52261	agcagaagcc	ctgaagatgc	agcttattct	ttagtcaagg	taacctttct	actaaatggt
52321	tcttaacaac	ttttaagtat	taaagtatta	cagtataact	ccaaagtatt	acaaaaggat
52381	cctgatggat	gaatatggag	aaaataacct	tagttccttt	gtgattcttg	gtgagtaagc
52441	atacctaagt	agcagagctt	attaattatt	caaaaatatt	tattatttgg	gtacaagcac
52501	tatgctaggc	atgagactat	aaatggagaa	agacagctta	gtcttgacga	cagagaccaa
52561	aaaaaaaaaa	atcacacaaa	tctataatta	cacattatgt	taagttatga	tggaaaaata
52621	tagatgctat	gacagtctac	atgagaggga	tctaattccag	tctggggagg	taagtgatgg
52681	ctaggctcag	aattcagatc	cgacctatgc	cacttactaa	cctggataag	ctacttacct
52741	ctctgaatct	gtttttctcat	ctgtaataat	agttcataag	agtatctact	tcataagggt
52801	cttataagta	ttaaattgaga	gaaatccatg	aaccagttgg	tatagtatct	agacactttc
52861	aatacagggt	ctgtatgttt	gattattttca	gggctgttcc	atgactcctc	cagtttgggt
52921	taactgcctg	tctctatgct	ccccaaaaac	gctctacata	ccactactac	aacattacct
52981	tactgtattg	caatcatcta	ctcatttaac	tattctgttt	tcattgagact	gtgagcagct
53041	ctgcaggaat	cacgccttat	gcttctttca	tccctaatagc	tccaagactt	ctagctcaca
53101	gctcgagctc	cttaaatggt	gtaggagaaa	taaatacact	cttccccagc	actcaccatc
53161	tttatttcac	tacagccttc	ttatcttttc	ttaacaagc	catctcagga	tctttacatt
53221	ggctattcct	tctgccatca	aatcaatcta	atcatctatc	tatctatcta	tctatctatc
53281	tatcWatcgc	aatcaatcaa	tcacaatttt	tctctattgc	aatcaatcaa	tcacaatttt
53341	tctctatcta	tctatctatt	gcaatcaatc	acaatttttc	cggctggggc	agggtggctca
53401	cacctgtaat	tccaacactt	tgggaggcca	aggcaggtgg	attacttgag	gccaggaggt
53461	cgagaccagc	catgccaata	tagtgaaacc	ctgtctctat	taaaaacaca	aaaattagcc
53521	gggctgtggt	gcggacgcct	gtaatcccgc	ctgaggcagg	aaaatcactt	gaacctggga
53581	ggtggagttt	gtagttagct	gagattgtct	cactgcactc	cagcctgggt	gagagtgaga
53641	ctctgccaca	aaaaaaaaaa	aaaaaaaaaa	gaacaaattt	tcgtatgctt	ttatagaaaa
53701	ttctttctcca	gatctccaga	tcttttcttt	tttgattgct	ccttgctggg	cagggctatc
53761	ccacaggcag	tatgcctaga	gtagccccag	atctttcaac	agcttgcatc	tcagcactct
53821	tgtctctgct	cagttacctc	tctctagagc	tgttttctga	ccaccttacc	caaaatagcc
53881	acttcacccc	aaactctgag	tcactctcta	accgggtcct	tattttatct	ctagcatgta
53941	ttacgtgaaa	ttgcattatg	atttttactca	ctctgtccca	ttagaatgta	agctaattag
54001	aattctgtcc	cactagaatt	ccaattctgt	cccattagag	tgtaagctcc	atgtggatag
54061	agtataaaat	taaaacaaac	tcagttttct	ctattatttc	actcaacaat	caacaacaca
54121	caacacttct	gataccagat	gtgtgatggt	ttttccccac	atatcaagca	agcttatcaac
54181	tgtgtctact	ggacaccagt	caagtgtcct	ctaattcaat	tcaattttga	tgttatctac
54241	ctggagatac	tgctcagatg	tacaggttga	aggctcagtc	ccacaagact	gcccttactt
54301	cagatgccaa	cagtaggtcc	aaggccactg	gccaactaac	tacaaaccag	ggttcccatg
54361	acctcatctt	caggttcaac	taatttgcta	gggcagctca	cagaactcag	ggaaatgtgt

54421	ttactgatatt	aaaggatatt	acaaagaata	ctgaggaaca	ccagatggta	ggataggaaa
54481	gctctggagt	cctcaagact	ggccatttaa	agatcgatt	ctgccactgc	aattttaagc
54541	acgcaatacc	catgacaggY	cagatgatgt	ggtggctcac	gccagtaatg	ccagcactgt
54601	gggaggctga	gatgggtaga	tcagatcagt	taagcccagt	tcaagaccag	gctgggcaac
54661	atggcaaaac	cccattctcta	ctaaaaaaa	caaaaattaa	ctgggcatca	tggcaagcgt
54721	ctagagtccc	agctactcta	ggcgctRag	gcaggaggat	cacctgagcc	ttgggaggtt
54781	gaggStgcag	tgggctgtga	ttgggccact	gcactccagc	ctgggcaaga	gtgagaccct
54841	atatcaaaaa	ataaaataaa	taaaaaatac	ccatgacatc	tctagaaatt	tttctgtgt
54901	gtccattttca	aattccacct	tttccacaaa	gtctcccctc	ataactttac	tagtctatga
54961	cacagtgtct	cacacctggg	gagcactcaa	atgtttgtct	cctgtaacat	ttactaaaca
55021	aactcttggc	tgggcatggt	agctcatgcc	tgcaatccca	gcactttggg	aggcctaggt
55081	gagaagattg	atggaggcca	aSagt tggag	accactctgg	ccaacgcagt	aacaccctat
55141	ttctagtttt	tttttaaaaa	aatttcagat	gccgggcatg	gtggctcatg	cctgtaatcc
55201	tagcactttc	ggaggccgag	acaggaggac	tgcttgaggc	caagaattca	agaccaacct
55261	ggccaacaca	gcgagaccgc	atctcaaaaa	ataaagttaa	aaaaaaaaagc	aaaaaaaaatt
55321	tagaaacgag	ggtggggaga	aaactctcca	gccacataa	cttattccat	tcaattcttt
55381	ttgcactaag	gttagtatta	cacaaatgca	cattttgtgt	tctgtaattg	ttttatgtct
55441	gtataccttg	cacctctatc	ttccctgttg	ccaaagacag	attttcttct	gtgtacacaa
55501	ccccaggac	agagcaaggc	accaacagta	attcccaaaa	taattttttg	attaaaactt
55561	ttggagtcat	actgtcatca	gccagaaaaa	aaaacatgcg	tgaggagaat	gaagggtgtt
55621	aaaagagctc	aagctgatgt	cctRaaagga	ttactttatc	tataaccttc	ctgaacacta
55681	gccttttctc	gttgagttcg	catgKaaagt	ttttccactt	aaaatggggc	agcaaggctg
55741	ggaactagta	ggagtagcgt	aacttctcgg	actctcccac	tgaaaagcat	cttttcttta
55801	aaggcggaga	gcactttaca	gacggttact	aaggttaaccg	Raaacagcaa	aactgagggtg
55861	gttctatcta	cggaagaaat	tctaagctcc	aaagttaaaa	tgcgagcag	tggcgcatat
55921	cagcagacgg	cggtcactac	tgtgtctcat	ttaaatctgt	gttcagctcc	gggctctttt
55981	tttcaggaat	ccttgtttcg	ggacacccag	tggacagcga	aaggagagag	accaccacag
56041	gcaaacggcc	ggaaagccgc	gcagtgcgag	cccgcgcgct	tactccgcc	caaataagcc
56101	actaccgtca	ccaaccccat	ctcgttcgcc	cgagccgcgg	acctgaggga	gcgctgtccg
56161	ctccccgcaa	gatgctcacg	cgaaaaaccc	actcttcagt	gtctgagacc	actccgcage
56221	ttccggcctc	ggtgagcggg	tgaactaaac	ccgggaaccc	gctgaaagac	ccagcttcgc
56281	actaggccta	gtcgtagcgg	cggccagcgg	gagggcgtag	ccgtagtctt	ggccacgctc
56341	attctgtgcc	ccgggtcagg	tcaccgcccg	gagggtatgg	gaagagaaat	gaagaagact
56401	ggtactccca	gacctttccg	tatcgaggac	cccaatcaac	aaccacctg	gtgagcgcca
56461	tcttcttccg	cttcttccag	cgcgagcgac	agcaccgctg	ggcggtctt	tccagatgat
56521	gaggagcgga	aaggaaggag	tgcgcgatgc	gcaggcgctc	tggccgaaaa	ggccttctgg
56581	gatttgtagt	tctttgaaca	ggagtaccgt	ttgtattctg	acSagaccaa	gcggttatca
56641	aggaataatc	tcttcagcta	gcaacaaatg	cataactcga	tcgcttatat	gcgacatgat
56701	agggcggttc	ctaggcaaaa	tatcaaagat	acagggagat	gccttttagg	ccgctttatt
56761	acgttctcta	ctcagtcccg	ttccccgaga	aggggtacaa	aatacaaaa	cttcccgggg
56821	acagMtRctt	taagtaagag	tgataggcg	tttaaaacga	caggtataaa	tgagactagc
56881	taatggagag	tYcttggaact	ctctgaaaac	attcacactc	aaaacttttc	tttaaacctt
56941	atcctctcac	ccgccaaaa	actctcttgg	ggagatattt	cgtaatcctt	ggtaattttac
57001	cgagtgtctg	ggttctctgc	ccagggtatg	acaaggctga	aatctcatag	gaggctctgg
57061	agaaaaatcc	atctcctaac	tcatctctgt	tggcaaaaat	cagttcctgg	cgggtgtggg
57121	tttcggcttc	ctgttttgtt	ttgtgctggt	agtgggtggc	accctcagct	cctagggggtc
57181	actctcgagt	ccttcccaca	tgagccctat	ttattggtag	ttcacacat	ggatatttgc
57241	tttctgaggc	acatctttct	gactcctctt	ctgccactgt	tggagaaaaa	tgcctgtctt
57301	ttaaagggct	catgtgatta	gatcaggagc	acctggataa	tctccctatt	ttaagatcaa
57361	ctcttggccg	ggcgagtag	ctcacacttg	taatcctagc	actttgggag	gtcggggcgg
57421	cggtattgcc	tgagctcagg	agttcgacaa	agcctgggca	acatggtgaa	acctgtctc
57481	tactaaaaat	acaaaaatta	gctgggcgtt	gtggcgcgct	cctgtagtcc	tagctactgg
57541	ggaggctgag	gcaggagaat	cgtgtgaatc	aggagacgga	ggttgacgtg	agccaagatc
57601	acgccactgc	actccagcct	ggcgacagag	tgagactgcg	tctcaaaaaa	taaaaaataaa
57661	taaaataaaa	aataaataaa	taataataaa	gaatagctgg	gtgtggtgat	gtgctcttgt
57721	tttctcagct	acttgggagg	ctgaggtaca	agaatcgctt	gaaccagga	ggtggaggtt
57781	gcagtgagcc	gagattgcac	cattgcactc	cagcgtgggt	gacaaagtaa	aattctgtct
57841	caaaaaaaaa	aaaaaaatca	actctgccac	accgcataac	agaatcagaa	tcttcatagt
57901	catggcccca	ggggttatgt	aggtcaggta	ctgtgggggt	tgggtgagga	ggagatcttg
57961	ggggatctcc	tagaattctg	cctgccacag	catgtaaacc	aaaaaaaaaa	aaagaaaaaa
58021	tcttttggtg	aatacaggat	tgttaatcac	attgtgtaat	ataacctgca	tggcctggga
58081	gcccgaagat	tagagttaa	gatgtagctt	tggctgggtg	tgggtggctt	tgcctttaat
58141	cccagcactt	tgggaggctg	aggcgggcag	atcacctgag	gtcgggaggt	tgagaccagc
58201	ctgaccaaca	tggagaaact	gcgtctctac	tgaataata	aaattagccg	ggcgtggtgg

58261	tgcttgccgtg	taatcccagc	tactcgggag	gccgaggcag	gagaatcgct	tgaactcggg
58321	aggcggaagt	tgtggtgagc	cgagatcggt	cgctcgttgc	ctctagccgt	agcaacaaga
58381	gtgaaactgt	ctcagaaaaa	aaaaaaaaaa	aaaaaaagat	gtagcttcag	caccaataaa
58441	ttcttatagg	gcaagtcatt	taattttatt	gggccccagt	tttattgaac	ttaaaatgct
58501	gatcatgtta	tctgctttac	ctatatttat	ataaatgtaa	aataattata	tacaatacag
58561	cagagtatga	gccttaaatg	aggtaatgta	ggtggatatg	tttttactta	actcttaagt
58621	actatgtaaa	tatgttagtt	attacagtta	atattttggt	tgacagtgat	ctacttcaac
58681	ctctgctcca	ggatgttatt	atttctaaaa	ggagaaaaac	agagaagcaa	attatagaac
58741	gtcctggaag	ccaagatcaa	ggagcttaga	tttagttcag	caggtgattt	gcagctgaag
58801	gaggacttaa	gtagagggcc	tgggtcaaa	gaacaggcaa	gaaagacaat	tcttaaggca
58861	gaatgggttt	ctatggcaat	gagagacaaa	ggaggctact	gaggacaagg	ttaaaataat
58921	taacatggga	aatcagtaag	acatagacca	agacctggga	tgaggattg	aaagaacagg
58981	gtggagaaga	gcttttgcaa	tccactctca	ttttccacag	aggcacatca	tgcaaggcat
59041	ttccaccagt	catagtccct	cgctgttttg	ggggtgtgta	aagaatagaa	actgacttta
59101	atagcaatcc	tatcaaaaac	attaagaaa	caaactctact	gcaaacgtgW	tacaatatta
59161	tacccaaagg	cttagtgctt	tctttgggtt	tggtggggag	gggggttggg	gttaagctat
59221	gctgaatcct	cagtgcctat	ccatccctca	ttatcatagg	taattagcag	ctgcctgctc
59281	atgtagagag	cccagtgcct	gctggaatca	gtaatgcgcc	aggttgcaag	aagcaatcaa
59341	ttcactccag	tcttagaatt	ctggatagag	agtagaattg	ttctcagaga	acctagggcc
59401	taatttgatt	cctaagtagg	tgtggctatg	tagccacaag	taatatagct	tgtgagggtg
59461	tagggaggac	aagttaagca	catattctgg	agtcagcttt	cctgccttct	gatcccgatt
59521	ccctaattta	tgagctgtag	attctggata	agtcacttct	ctctgtctca	gttttctcac
59581	tggccaataa	aaagggtgag	ggcggtggaa	tcagtataac	ttatcctcaa	taaagctatt
59641	gggaggatta	aatgagttaa	tcatgtgaag	tgcttagaaa	aatgcccagc	atagagtaag
59701	cgctcaaaaa	ttaattatgc	tttctggctt	taatgaattg	gccttaattt	attaaggttc
59761	aggtttttga	tctgcaaagt	gaggcagttg	aactcaattt	tatttctggg	gtaccttccc
59821	actaacattt	tctactatag	gatgtttttt	taaatttact	tcaatgccag	tattttatgc
59881	tctgatgcca	ttatgcttta	gttattatgg	ctatattatt	ataatcattc	tctgataaaa
59941	agtagagctg	ctccaccWca	ccgttctttc	tttttgatg	tgttttacgt	gccctctcaa
60001	tgtttaaaat	ttttttttta	ttaaaaatatt	ggtttgcatt	aaagcatagt	taaaagggtc
60061	agagccaatt	attacttagg	aactgcgaaa	gaaataacgt	atctttacga	tggagagatt
60121	tggtgtgcat	cacttttgct	tcactaacag	tggacaacatc	tgacactagg	tgcttttaat
60181	aggaagtccc	ccagttttct	ttctttcttt	ctttctttct	ttctttcttt	Yttttttttt
60241	tagaccgagt	cttcctccgt	tgcccagggt	ggagtgcagt	ggtgcgattt	cggcttactg
60301	caacctctgc	ctccctagtt	caagcgattc	tcctgcctca	gcctcccag	tagctgggat
60361	tacaggcagc	agcaagcacg	cctggctaat	ttttgcattt	ttagttaga	tggggtttca
60421	ccatgttggt	caggctagtc	ttgaactcct	ggcctcaagt	gatccacca	ccttggcatc
60481	ccaaagtgtc	gggattacgg	gtatgagcca	ctgcacctgg	ccgRaagt	accaattttc
60541	tgatcaaacc	agacctaatg	tccagttcac	aagaaattca	ggggatagag	taacaagttt
60601	aaaagcatga	ggaacaatc	agtaaaatcc	agaatgtgag	acattctaca	agacaactga
60661	ctgagactct	acaaaatgtc	aatgtcatgg	aatggaatg	aatagaatag	atgaaaacag
60721	attgatgaga	catgacacat	aaaattagta	aaccttgatg	ggattctggt	tcagttgggg
60781	aaaactataa	aaaatattct	tgggacattt	ggagaaagtt	gcataatgg	tgataataga
60841	caacattgta	gaatttttgc	ttaaatttctt	aggagtatta	atgacattgt	ggttaagtag
60901	agaacgctt	ttattcttgg	gagatgcatt	ccagaggatt	tggggatgaa	atgtgtcaac
60961	ttacttttct	cttttttgtt	tgttttttgt	tttgagacac	aattttgtct	ttgtcaccca
61021	ggctggagta	cagtgcgcga	gtcttggtt	actgcaatct	ctgcctccc	ggttcaagt
61081	attctcctgc	ctcagtcctc	caagtagctg	ggattatggg	cgctgccac	catgccagc
61141	taatttttgc	gttttttagta	cagacgaggt	tttaccatgt	tggccaggct	ggtctcgaa
61201	tcctgacctc	agggtgctca	tccacctggc	ctcccaaagt	gctgggatta	taggcctgag
61261	ccactatgcc	cagcctcaat	ttactttcaa	atagtttatc	tcaacaacaa	aaaatgtgtg
61321	tgtgtgtgtg	tttatagtga	gagagtgtga	gagagcatga	aagggaaagc	aaaatgttaa
61381	caattatttta	agcttagatg	ggggagtgtg	gactgtatag	gtgtttattg	aactattgtt
61441	tcaactttta	tatatatatg	aaaattatct	cttttttttt	tttgagacag	ggtctcactc
61501	tgttgctcaa	gctggagttc	agtggtagta	tcacggctca	ctacagccc	aatctcccag
61561	gctcaagcga	tactcctacc	tcagcctccc	aagttagctag	gactacaggc	atgtgccacc
61621	aggcctggct	aattttattt	ttgtagagac	agggtctcac	tatgttgccc	aggctggtct
61681	tgaactccc	gactgaagt	atcctcctgc	ctggcatccc	aaagtgatgt	aagccactgc
61741	accagccct	gaaaattttt	ataataaaaa	tttaagaaga	ggaaagtgt	agaactcaca
61801	tttcccaatt	tcagaaacct	actacaagc	tcagtaatc	aagacagtgt	ggtactggca
61861	ttaggataga	Ytcatagatc	aatggaataa	aattgagagt	cctgaaataa	accatattgt
61921	ttatgggtgaa	ctgatttttg	acaaggttac	caagaccatc	taaaactatt	attgacatag
61981	ttttttcaac	aaatggtgct	gggaaactgt	atgtctaaat	gcaaaagaat	gcagttggac
62041	ccttgcttca	cgccctatac	aaaaattaac	tcaaaaatac	tcaaaagacct	aaatgtgaaa

62101	actaaaacta	taaaactcct	agaagaaaac	acaggggtca	atcttcatga	tcttggattt
62161	ggcaaagaat	tcttagataa	gatacaaaaa	gcataaacia	caaaggaaaa	aatagaaaaa
62221	ttggaYtaaa	tcaaaagtta	aggttttgtg	cttcaaagga	catcatcaag	aaagtgaagg
62281	gccgggcatg	gtggctcacg	cctgtaatcc	cagcactttg	ggaggccgag	gcaggcggat
62341	cacgaggtca	ggagatcgag	acgatcctgg	ctaacatggt	gaaaccccat	ctctactaaa
62401	aatacaaaaa	ttagccgggt	gtagtggcgg	gcacctgtag	tcccagctac	tcagaaggct
62461	gaggcaggag	agtggcgtga	acccaggagg	cagaacttgt	agttagccaa	gatttgtcca
62521	ctgcactcca	gcctgggcaa	cagagcgaga	ctccgtctaa	aaaaaaaaaa	aaaaaaaaag
62581	aaagaaagtg	aaaagtccca	agaccttcag	gcaaaciaag	acatttcctat	taggcatgac
62641	atttcaggag	tttagagatt	acttccaaga	agctaaggac	aaaagccaga	tctctctttg
62701	gacaaggtta	aatttgttac	tacagaagcc	ttgaaaacat	tatgctaaat	gaaagaagcc
62761	tgtcacaaaa	gaccagatat	tgtattatta	cctttatata	aaatgttcag	aataacaaaa
62821	tctgtggggg	gtgttaggaa	aggtaataaa	ttgttaatgg	gtatgagttt	cttttggggg
62881	aggggttgaaa	aatttagtgg	gacggttggt	caactttgaa	tatactaaac	gcctctaate
62941	tgcatacttt	aataagggtga	ttttatggtg	cgtgaactac	atltcaataa	agctgtagaa
63001	gcaaagttaa	gtaaaatgca	aaagtaataa	aatgaataag	tgaataaaaa	gtataatata
63061	gtatctcttc	atltttttcc	aaatgaaatt	tagaattatt	ttgtgcattc	tattgaaaaa
63121	tatctgtagg	cattttgatr	ggcgttgaat	tgtccataaa	ttatltttga	gagagccaat
63181	atctttacaa	ttttatgtct	tctcactcaa	gaacactggt	tttggttttK	gtttttgagt
63241	tggagtcttg	ctctgttgcc	tgggtctggag	tgcagtggac	gatcttggct	ccctacaacc
63301	tctgcctcct	gggttcaaga	gatttctctg	cctcagcctc	ctgagttagct	gagattacgg
63361	gcatgcacca	tcacacctgg	ctaatttttg	tatttttagt	agagacgggg	tttcgccatg
63421	ttggccaggc	tgtctctgaa	cccctgaact	caagtgatct	gcctgccKtg	ggctcccaaa
63481	ctgttgggat	tacaggcggt	agccaactgct	tttggccagg	aacactgttt	agattttaac
63541	ttattaaaaa	cttccttcat	atctttcaat	aacattttga	agtgtgtctt	atatagattg
63601	tacacatctt	tcaggatatt	aattttgtga	tttcattttt	aaaatcatga	atgcaatctc
63661	tgtttttata	tgattactaa	atagaacagt	aattttttatt	aagtaagtca	aaaatgttct
63721	gaaacatgtc	agtcatttta	taggcagaat	caattttttt	aattttttatt	ttttattttt
63781	tacaggYatg	atcaaccagg	tacagaatca	gttttttggc	tgtagatgat	ttaaagagta
63841	gggtcaactt	gtctttgggt	agatgtgtct	ctctatcttc	ttctggaatt	aagcagttct
63901	ttgggttgat	gagttaggaa	gacttcttgc	tgtcagctac	ttccctccat	tgctcttgag
63961	acttcctggt	taaaagcact	acgaaacctg	ctgacccccc	ttctttcccg	atttaattat
64021	tctttttgaa	agttggctct	aaccttaagg	gttgtgaaaa	cacaaggatc	cagtgtgggt
64081	gaagcttttag	aactaagaaa	accatttcctt	tttctatcaa	acagcctggc	aggggaacag
64141	tttgccccct	gccctgagcc	ctgtcatttg	ttttgatcat	tttcatctct	atcaaatgtg
64201	ctgttggtccc	agaaaggcagc	catctgatag	ctgaagtgcac	tgcctacatt	tcttccctct
64261	cttctccagt	aacaagatat	tccaccagaa	tatgcctggt	tgtaggctaa	ggctgtgatt
64321	taataaccct	aaggagataa	acactgggat	aataataatg	gttattattt	attgattagc
64381	ttctctatgc	caattttatcc	cacttatttcc	ttgtaacagc	cctgtgagta	gttaactctt
64441	tttatccaca	ctttacagat	aagaaaattgg	aggcttagat	atgtaagctg	cccaagatga
64501	caaataagaaa	gtgactaatc	tggaaacttaa	acccagagaa	catgatcttt	cttttttttt
64561	ttttgagaca	gggtctcact	ctatcgcccc	ggctgggggtg	caggggtgca	atcacggctc
64621	attgcagcct	tgacttcctg	ggctcagtct	atcctcctgc	ctcagcctct	caggtagctg
64681	ggactacagg	cgaggcccac	cacgcccagc	taattttctt	atltttttgca	gagatggggg
64741	cccactattt	tgcccagggt	ggtcttgaac	tgctgggctc	aagcaatcct	cctgcctcag
64801	cctcccaagg	tgctgggatt	acaggtgtga	gccactgtgc	cgggcccggg	gaacatgatc
64861	ttaaccatac	tgtcttccaa	gtgaaagatt	ctacctttct	aagacatctg	gagtcagaca
64921	gacctgagtt	caagtctgtc	cttcctcccc	actttactga	accttcagtg	agtctttcta
64981	cctcagaagg	ctgtgaggat	gccatggggg	aatgaatgta	aagtgtttaa	gtgcttggct
65041	gacactgagt	actaaataaa	tatcaattta	ttcttcaaaa	gtgtataaaa	atgtatcaag
65101	tcttttctct	gtactaaacc	ctgtactaat	ttgggggaat	tcaaagttaa	agctgtgata
65161	gtcttgccct	aaaaactaat	gggtctgggac	gggcacagtg	cttcacacct	gtaatcccag
65221	cactttggga	ggctgagggtg	ggcagatcat	gagatcagga	gttcgagacc	agcctggcta
65281	acatatgaaa	ccccatctct	actaaaaata	caaaaaatta	gctgggcatg	gtggcggggtg
65341	cctgtaatcc	cagctactcg	ggaggctgag	gcaggagaat	tgcttgaacc	cgggaggcag
65401	agggttcagt	gagctgagat	cagcccatgtg	cactccagcc	tgggcaacaa	gagcggaact
65461	ctccatctca	aaaaacaaaa	acaaaaatac	aaaaccaatg	ggctgggtaca	tctgggactt
65521	gtgggttgaa	tctggccccc	atacatattc	tgtttgtcct	tcacatgctt	taaaactttt
65581	gaattagttc	cagacattta	aaagttggga	gattttggcc	gggcatgggtc	gctcacacct
65641	ctaactctag	cactttggga	ggccgaggcc	agtggatcac	ttgaggtcag	gagttcgaga
65701	ccagcgtggg	caacatgggtg	aagccccatc	tMtactaaaa	atacaaaaact	tagccaggca
65761	tgggtggcggc	acctgtaatc	ccagctacat	gggaggctga	ggcaggagaa	tcgcttgaat
65821	ccaggagggtg	gaggttgcaag	cgaagctgaga	tcttaccact	gtactccagc	ctgggcgaca
65881	gagttagact	ctatctcaaa	aaaataaaaat	ataatataaa	ataagttggg	agattttgat

65941	ctgtttcacct	gcctttgaac	ttcctaccca	gacaccttct	caagggggRt	ggtgtctctg
66001	gcgatctcat	cactggcaga	ggaaggataa	ctgaaccag	tcaatggcca	agcctccata
66061	tattagtggc	actcacatct	ctagactctt	cttggtcac	cagacataca	tgtaccttgg
66121	ggaccaccct	tcaaaatgcc	atcaacacag	gctcctaaag	taggacattt	tacacccatc
66181	accagttat	gtaccaacaa	taacataggg	ctattccttc	agatgaagag	ctgagaatat
66241	gttagtttta	aaatagaagt	cagagtcagt	gtggcaggtg	gaattttacag	cctgtgtaca
66301	aaaaaaacca	cacttaaatg	cctgtttcta	acaaagacca	actaaaaatt	cctcccaggg
66361	caataaagag	aagcctccat	agggtgaggt	gtttctctgg	agggaatcta	gaaaagttga
66421	gggagaatga	gtctagctcc	ttccaacagg	aggatagaaa	cctatctaaa	tgggcagggc
66481	cagactgagt	tcataggaaa	aataatttcc	ttcatagtgg	gtagattcta	agacccccac
66541	cccaattttt	tttgaggata	tctcacaagg	gatagagcag	gcctgggtct	gacaggtggg
66601	actgacttaa	gcctatcctt	gctcacaac	agaagcccat	tcccaccagc	gatctgagag
66661	atttggcaga	ggctgggtga	aggttgagga	tgaaggggac	tttacttgtg	gagttcaagg
66721	ttaaatccaa	tctacagtca	tgaactctgc	tccaaaccac	aataccgctt	cttaccact
66781	aggatagcta	taaccaaata	gacataataa	gagttagtga	ggatgcagag	aaactggaat
66841	gcttgtacat	tgcaggtggg	aatgtaaaat	ggtgcagcca	ctttggaaaa	cagttttggca
66901	gttcttcaaa	actttaaacg	gagagttacc	atatgatatg	gtttggctgt	gtgtccccac
66961	ccaaatctca	tgttgaattg	taattcccaa	tgttggggga	gggacctgat	gggaggtgat
67021	tggatgacgg	gggcagattt	cccccttgct	gttctatgat	agtgagttct	gatgtcctat
67081	gatagtgagt	gagatctgat	ggtttaaaag	tgtgtggcac	ttcccccttt	gctctttctt
67141	ctgccaccat	gtgaagaagg	tgtcttgctt	cccctttgcc	ttctgccatg	attgttaaat
67201	tcctgaagcc	tcccagtcac	gcctctgtta	aacctgcaga	actgtgagtc	aattaaacct
67261	cataccttag	tacaccatat	gacccagcaa	ttccattcct	tttttttttt	tttttttttt
67321	ttttttttga	gacaggtgct	aactctgtca	cccaggctgg	agtgcagtg	tgcaatcaca
67381	gctcattgct	gcctcaacct	cctgtgtctg	agtcacctc	ccacctcagc	cttctgagta
67441	gctgggacta	caggcccatg	ccaccatgcc	tggctatttt	ttttttttta	tttttggtaa
67501	agacaggatc	tccctatggt	ttccaggctg	gtctcaaaact	cctggactca	agcaatcttt
67561	cagcctctat	ctcccaaagt	actgaggggg	attataggtg	tgagccactg	ctcccagcct
67621	ctactcttaa	tatgatcaag	agaaatgaaa	acacatgttc	acacaaaaac	ttatacgcaa
67681	atattcatag	cagcattata	atagctgaaa	tgtggaagca	gctcaaatgc	gccacacact
67741	gtgaatgga	taaacaaaag	gtggtataatc	cacacattga	aatattactc	gacaataaga
67801	agaaatgaag	tactgatacg	ggatacaact	tacacgaacc	ttgaaaacat	aatgctaagt
67861	gaaagaagcc	agtcacaaa	accagatatt	atataattcc	atttatttga	attgttcaca
67921	ataggcatat	ccatcggaat	gcagattagc	agttgcctgg	ggatagggcc	ggtgaggtgg
67981	ggcaggaagc	agggactgag	aaagggaaca	aacagcgagt	ggatgggagt	ttcttttttg
68041	agtattgaaa	ggtaaaaagt	tctaaaatta	tataatggtg	atgggtgtac	aactctgtga
68101	ataatcatga	aaaccactga	attgtacaca	ttgaaaggg	gaattttacg	gtatgtgaat
68161	tatatcttaa	tacatctgtg	ttttggccag	gcgtggtggc	gtggtggctc	acacctgtaa
68221	tcccagcact	ttggggaggc	gaggcaggcg	gatcacttga	ggtcaggagt	ttgagaccag
68281	cctgatcaac	atggagaaa	cccgtctcta	ctaaaaatac	aaaattagcc	aggcatgggt
68341	gcacacacct	gtaatcccag	ctactcggga	ggctgaggca	agggaattgc	ttgaacccag
68401	gaggcagagg	ttgtgggtgag	ctgagatcgt	gccattgcat	tccagcctgg	gcaacaagag
68461	caaaactctg	tctcaaaaaa	aaaaaaaaat	ctgtttttta	aaaatatatc	tgttccaggt
68521	ttcctcagac	tacccttaag	ataaccagga	accacaatgg	acttttccct	ctttctcctc
68581	tatttttaaaa	aaatattcag	ccaggcatgg	tgttgacact	gtaatcccag	ccacgtggga
68641	ggctgagttg	ggagaattgc	ttgaacctgg	gaggtggagg	ctgcagtgag	tcgagatcgc
68701	accactgcac	tccagcctaa	gcaacagagc	gagactccat	ctaaacaaca	acaaaaattc
68761	atgagcctct	ctttacttaa	tttctttaaa	atgtttaata	gtctgttaat	taattttataa
68821	aattattatt	aKtagtattt	gagacagagt	ctcactctat	tgccaggctg	gagcacagt
68881	gtgtgatctt	ggctcactgc	aacctctgac	tcctgggttc	aagtgattct	catgccttag
68941	cctcctgaat	agctgggatt	acaggcatgc	accacaatgc	ccggctaatt	tttgtgtttt
69001	tagcagagat	ggggtttcat	catgttggcc	aggctggtct	cgaactcctg	tcctcaagt
69061	atctgcctgc	ctcagtcctc	aaaagtgtct	ggattacagg	aatgagccac	agtgcctggc
69121	caaaaaaatt	tttttttctt	tttttttttt	ttcagagaca	gggtctcact	gtgtaacca
69181	ggctaaagt	cagtgcaccc	atcatgactc	actgcagcct	cgatttcttg	ggctcaagt
69241	atcctcccac	ctcagcctcc	tagataactg	ggaccacagg	cacgtaccac	catgccagc
69301	taggtgtatt	agtctgttct	catgctgcta	ataaagacat	atccgagact	gggtcattta
69361	taaaggaaag	aggtttaatg	gactcacagt	tccacatggc	tgggtgagacc	tcacaatcat
69421	ggtggaaggt	gaatgaggca	caaagtcacg	tcttcatatg	tggcaggcaa	gagtggtgtc
69481	aggggaattc	ccgtttataa	aaccatcaga	gcttgtgaaa	cttatttact	atcatgagaa
69541	caggcagagg	aagaactgcc	cccatgattc	aattacctcc	cacgggttcc	ctccccagc
69601	atgtgggaat	tatgggagct	acaattcaag	atgagatttg	ggtgaggaca	cagccaaacc
69661	atatcactag	ggtttttttt	ctttgataga	gatggagtct	cactatgttg	cccaggctgg
69721	tcttgaactc	ctgggctcaa	gcaattctcc	cactttggat	tcccaaagtt	ctgggattac

69781	aagtgtgaac	cattgcgcc	ggcctcatgg	cctagttaac	tgcttctttt	ttttcccgag
69841	atagagtctt	gttctgtcac	ccaggctgga	gtgcagtgg	gagactcagc	tcactgcaac
69901	ctcYgcctcc	tgggttcaag	cgattctcct	gcctcagcct	ccagagtaag	ctgggattac
69961	aggcaccgcg	cactgcacct	ggctcatttt	tgtattttta	gtagagatgg	ggtttcacca
70021	ggttgggchg	gctgggtctg	aactcctgac	ctcaggtgat	ccccctgccc	tctgcctccc
70081	aaagtgcctg	gattacaggc	gtgagccacc	atgcctggct	agttaacttc	tggacctcaa
70141	tttcctcagt	tgaaaaatga	ggataataac	agtcctttcc	tcatagggta	agtgtaaatc
70201	taaaatgacc	acagatatga	actatgttta	gaacagtacc	tagcacataa	taaatcctgt
70261	attagtagtt	gtaataatta	tttatatatt	aatatatgtt	atacatatta	atagtcatat
70321	attacacccc	taggatagta	tttagcacat	agtaagtgcc	ctatRtaagt	gttagctatt
70381	aattattatca	ttatttactt	atttttatta	gagttgataa	tggttaactt	gtccaaggct
70441	ttatctgttc	tgctgttgat	gaacattgag	ctgtttccag	tttgagggtta	caaatgcgta
70501	atgctgcttt	gaatatgtgt	gcattcatttt	ctctagggtta	tacacttagg	agtggattgc
70561	tgaggcacag	ggtatgttta	cctccatatt	tacagatatg	ccaaaccgtt	ttccaaagtg
70621	gatgattcaa	tttacattcc	cacctccagt	aagaacctct	ccaatacttg	ataggctcat
70681	ctttttaaatt	ttaccctgg	tatgtaatga	cagcatcaat	tttaatttgc	atttttctgg
70741	ttactgatga	ggaggagcac	ctttgatttc	cttctgtgta	cattcctgtt	caagcctctt
70801	gtccattttac	tactgagtta	tacgtctttt	tctgtgtgcg	ttgtgagaat	tcttcacatg
70861	gtttgggtac	tgagtaattc	tttatNtatt	tcaatacctt	ccactattct	gcttgacttc
70921	tactctgcca	gtggtatctt	ttgtttttgt	ttttgttttg	agatggagtt	ttgtctttgt
70981	tgcccaggct	ggagtgcatt	ggtgctatct	cagttcactg	caacctctgc	cttccattgt
71041	caagcgattc	tctgcctca	gctctcaag	tagctgggat	tctaggtgcc	tgccaacacg
71101	ccggggcaaat	tttttgtatt	tttagtagag	acgggggttc	accatgttgg	ccaggctggg
71161	ctcgaactcc	agaccttaaa	caatccaccc	gccttggcct	cccaaattgc	tggtgattaca
71221	ggtgtgagcc	accatgcctg	gccccagtg	tatcttttga	tgaacaaaaa	aattttaattc
71281	agaagtaagc	aaatttatga	atctttttca	ttgtagttag	tgctcttttt	gtcttgttca
71341	agaaattctt	ccctgtcctg	aggtcttata	ttgtaaaagc	tttatagttt	ttctttttac
71401	acttaggtca	tttccacctg	gaattgaata	ttgtgcacga	tctgatatag	gactgagggc
71461	aatttcaatt	ttttctacat	agatgctcaa	ttttatcagc	accatttatt	aaaaagcctg
71521	ttctttcttc	actcatctgc	aggccatctc	tgtcatgtag	caagtgttat	gtagcaaatg
71581	tcatgtattg	atatgtttga	gggctcttta	ttttgttctt	tttacaggaa	gtctctccac
71641	cttgtttctc	ttattcaata	gtgtttttgt	ttgatccctt	ttagaatctg	ctaattaagt
71701	tccaccaaaa	acaaacaaag	aaacaaacaa	aaccctaaaa	tggaacaaaa	ccattttccc
71761	agtcagtaga	aatctcaaga	tctaactgaa	tggtagtgc	caggactctt	tcttttttca
71821	cctggacatg	aacgaaaagc	atgtagattt	aggttctgtc	aacagctgcc	tcattctagg
71881	cacattagta	tagggttgaa	aacaacactc	acagccgggt	gcaggggctc	atgcctgtaa
71941	tcccagcact	ttgggaggct	gaagcagggt	gatcacctga	ggtcaggagt	tccagaccag
72001	cctgacaaac	atgggtgaaac	cccgtctcta	ctaaaaatac	aaaaattagc	tggtgcatagt
72061	ggctaacacc	tgtaatccca	gctacttggg	aggctgagga	aggagaattg	cttgaacctg
72121	ggaggcggag	gttgacgtga	accgagatcg	ggccattgca	ctccaggctg	ggtgtcaaga
72181	gtgaaactcc	atcacacacc	acatacacac	acacacacac	acacacacac	acacacacac
72241	acacacacac	aggctgggag	cagtggctca	cgctgttaat	cccagcactt	tggtgaggccg
72301	agggtgggag	atcacgaggt	caggagttcg	agaccagcct	ggccaacata	gtgaaacccc
72361	atctctacta	aaaatacaaa	aaattagctg	caagtgggtg	tggtgcgctg	taattccagc
72421	tactcgggag	actgaggcag	gagaatagtg	tgaacccagg	aggcagagtt	tgcatgtgag
72481	cgagattgtg	ccattgcact	ccagcctggg	cgacaagagt	gaaattccat	ctcaaaaaaa
72541	aaaaagaaaa	caacactcag	aagggcagag	ccaagagaat	cactaagaac	agagtcagac
72601	aatgtctggg	tccttctttac	ctcctaagtg	ctgcagttat	atcggtttaa	cttgggggtt
72661	ctgttatattg	caaaccactg	aaagcatcct	aacttatcca	taagctatct	taaatactgg
72721	cttggctgca	gaagcaagta	aaaagacctc	ttcgtgaaaa	ggaaaaatac	tgtcacaaaa
72781	ctattaattt	tattctatgg	gggccaacaa	tgatataata	atataataat	ggtgatgttg
72841	atagttatag	ataacaatta	caaagccagt	ccctatgcta	aggatattac	agataatata
72901	tcattatact	agtgatctta	ttcaacggcc	atgatctccc	agtgttactc	tttttagaga
72961	ggaggaaaca	ggtccagaga	tcttaagtaa	ctctctcagt	atcacacagt	taatggaatt
73021	caaaggtagc	cttgactggc	tccagcacac	ttatcaattg	gtcttttaac	taattattat
73081	tattattWtt	ttttttttga	gaYggagctc	cgctctgttg	cccagggttg	agtgacgtgg
73141	cacgatcttg	gctcactgca	agctccgctc	cctgggttaa	cgccattctc	ctgcctcagc
73201	ctcctgagta	gctgggacta	caggcatctg	ccaccatgcc	cggctaattt	tttgtatttt
73261	tagtaaggac	agggtttcac	cctggttagcc	aggagatggg	ctccatctcc	tgacctcatg
73321	atccaccctc	ctcggcctcc	caaagtgtcg	ggattacagg	cgtgagccac	cgcgcccggc
73381	cttaacttat	tattaaatag	agcaaatgta	aaacttccag	gaggccaggt	gtggctcact
73441	atgcctataa	tctcaacacc	ttgggaggcc	aaggcagggt	gatcacctga	ggttggaagt
73501	ttgagaccag	cctggcgaaac	atggtgaaac	cctctgtgta	ctaaaaataa	aaaattagcc
73561	aggtgtgggt	gcacctgcct	gtaatccag	ctacttggga	ggctgagaca	ctagaatgac

73621	ttgaaccag	gaggcagtg	ttgcaatgag	ctgagatagt	gctactgcac	tcagcccg
73681	gagctgtttc	aaaaaaacaa	taacaaagac	aaaaacaaaa	acttccagga	caggcggtgt
73741	gggtcatgtc	tttaattccta	gcactttggg	aggctgaggc	aggaggtttg	tttgagccca
73801	ggagttcaaa	actagcctgg	tcaacatagt	gagaccccat	ctctataaac	aacgtttctc
73861	ccatgtctgt	ctcatgatag	tgagtgaagt	ctcacaagat	ctgatggttt	tataaggggt
73921	tcttccccct	tcacttggca	cttctccttc	ctgccacctt	atgaagaagg	tgctggcctt
73981	cccccttgcc	ttctgccatg	attataagtt	ttctgaggcc	accccgcca	tgctgaactg
74041	tgagtcaatg	aaacctcttt	cctttataaa	ttaccagta	atttataaca	acaactcagg
74101	cagttgttga	tagcagcatg	aaattggagt	aatacagcct	cctaaagtgc	tgggattaca
74161	ggcatgagcc	accacacgca	gcctgcattt	tttcatagat	aaaatgttct	acaagttttc
74221	gattgacgtt	tggaaatttt	agaaaagtac	aagaagaaaa	tttaaattat	ccataatcct
74281	actctgcaga	gataatcact	attatcattt	tatagtatat	cctttgagcc	ttttaataaa
74341	tatatgcaat	agaaatccag	ctgtgaatat	tgttttgtaa	tctgctttct	tttttttttt
74401	tgatatggag	tctatctgtc	actcaagctg	gagtgacgtg	tctcaatctc	agcttactgc
74461	agcctccacc	tcccagggtt	aagcgattct	cctgcctcag	ccttccaagt	agctgggatt
74521	ataggcgctt	accaccacac	ctggctaatt	tttgatttat	tggtgcagac	aagtttctgc
74581	catggtggcc	aggctgggtt	cgaactcctg	acctcaagtg	atctgcccac	tttggcctcc
74641	caaagtgtct	ggattacagg	tgtgagccac	tgacactggc	ctgctttctt	aagatagtat
74701	tttttctagt	cattaaaaat	tctaaaaat	tattatataat	gtttgcaatg	gtatttttagt
74761	ttctaacaat	tacttaatac	gtgctctatt	ttcactaaaa	atttggtggc	atacacattc
74821	ttgaaaaatct	ttttttcaga	caacactgat	ttttttaaaa	gtagaattac	tttataaagg
74881	aatataaaca	tttacaagtg	tttagtacta	ccaaactgct	ctccagaaag	gctacatcca
74941	cttctgttcc	taccattagt	gaggtagaat	tcctcttttc	ccaaaacctt	actcttctgg
75001	aagttttttt	tttttttttt	taaatagaca	tggtcttgct	ctgtcaccca	ggctggagtg
75061	cagtgggtct	atctcagcta	actgcagctt	tgacctcttg	ggctcaagcg	atcctcccac
75121	atcctaccct	agcctcctag	gtagctggaa	ctacaggtgc	actccaccat	gcttggctag
75181	ttttttcttt	gtttgtttgt	tgagcccat	gattcaaata	cttccacctg	gtctctccct
75241	tgacacctgg	agattatggg	gattatgggg	attacaattc	aagggtgagat	ttgggtgggg
75301	acacaaagcc	taaccatatt	agaggccaag	gtgggaggat	tgcttgagcc	taggagtaca
75361	agaccagcct	gagcaacaga	gcaagattgt	ccacaaaaaa	aggaaaaaga	atttaaaaaa
75421	gacaaagaaa	acaaaaaata	aaaaaagtta	ttagatgaat	gtgggtggcc	atgctgttag
75481	ttccagctac	ttgggaggct	gaggtgggag	gatcacttga	gcctgggagg	tcaaggctgc
75541	agtgagctat	gattgcacca	ctgactcca	gcctgagcaa	gagagcaaga	ccctctcaaa
75601	aaaaaaaaaa	aaaaaaagaa	agaaaattca	aacattttca	aagctagaga	aaagagaaaa
75661	atgagatgaa	cttgccccag	ctttaattat	tgtcaacaca	aagccaatct	tgctgtcat
75721	attccttctc	cctcactcac	attattttaa	agcaaataac	aaacatcata	tttcatttgt
75781	aaatatattt	gcatgtttct	ctaaccaata	agggctcctt	tttattttta	tatttttttg
75841	agacagggtt	tcactcttgc	ttagggtgga	gtatagtggg	gtgatcatag	ctcactgcag
75901	ccttgaactc	ctggggtcaa	gtaatcctcc	tgcYttggcc	tcacaagtag	ctgggactac
75961	aggctYgtgc	caccacgtct	ggctaatttt	taattttttg	ttgagatagg	atcttgcatt
76021	gttgcccacg	ctggtctcaa	actcctgagc	tcaagcagtc	ctcctgcctt	gacctcccaa
76081	cgtgttagga	ttacaggcat	gagccactgc	acctggccaa	gggctccttt	ttttaaaaaa
76141	cacaaccaca	gtaccatgat	cacatctaaa	tttaaaatag	ttaattcttt	aatatcatca
76201	aatatccaaa	taatatccac	attgccctga	ctgtgttttt	atgtgtttta	tatatattat
76261	ctcttttact	taYaggctcc	ccctccatct	cttttttctc	atgtaatatt	ttgttgaaga
76321	aattgggtca	tttgccctgtt	agtttgccat	agccttggtt	ttgttaattg	tagctcatga
76381	tgttgtttta	catgtccctc	atctgtctgt	ttcctgtaaa	ttcttcatta	gacctagagg
76441	cttgatcaga	ctcagagttt	ttgttttgtt	ttgtgcagta	cagtgggaga	agggacgaag
76501	aaacttatca	aggggaggca	ttactttcta	tctgaaggat	ccagtgtctg	atggtctttc
76561	ttgtgatgtt	agcaactggt	tatgatcatt	gccacaccca	tacatcatta	gggatgtcaa
76621	attggcgatg	ttcttattct	attgtttctt	ttccatttag	tagttagaat	gctccatatt
76681	ccaagagaat	tttttcttta	actattttag	tgctcttttg	aagagatcac	aaaggaaaga
76741	gaaaataaat	acctgattct	ttctctttta	tgaccagttt	taaaaataat	taggtgcggc
76801	cgggtgcggt	ggctcacgcc	tgtaatccca	gcactttggg	aagccgaggc	gggcggtatca
76861	tcaggtcagg	agatcgagac	catccgggct	aaaacggtga	aacccgtctt	ctactaaaaa
76921	gacaaaaaat	tagccggcg	tagtggcggg	cgctgttagt	cccagctact	tgggaggctg
76981	aggcaggaga	atggcggtgaa	cccgggaggc	ggagcttgca	gtgagcggag	atcccgccac
77041	tgactccag	cctgggagac	agagcgagac	tccgtctcaa	aaaaaaaaaa	aaaaaaaaaa
77101	aaaaaaaaaa	aaaaaaatta	ggtgcttttc	tagcatgctt	cagtggtgac	tgataaaatt
77161	tttttggtaa	cgttataaac	tcacagagtt	aaataaactt	aatatgcttt	catccagtga
77221	tgtttttagc	ctattgatgc	tcaaatatct	ctccttttag	caagtagaag	gttctccatg
77281	ttagcttctg	agtccttttt	aaaccaaccc	cattagtctt	tgatggcttc	cctgctttct
77341	ggtataaaa	gtattccagg	ctggggggcg	tgactcatgc	ctgtaatccc	agcaatttgg
77401	gaggccgagg	ctggtggatc	acctgaggtc	aggagtttga	gaccagcctg	gccaacatgg

77461	tgaaccccg	tccctgctaa	aaatacaaaa	aaattagctg	ggcctggtg	tgggcggtc
77521	taatcccag	tacttgggag	gctgaggcag	gagaattgct	tgaacccag	aggtgacagt
77581	gagctgagat	tgcaccattg	cactccagcc	tggccaacaa	gagcaaaact	ccatctcaaa
77641	aaaaaaaaa	aaaaggtatt	ccaggtatcat	gttgtagatt	tccatcccca	gatctggaat
77701	gactcatttc	tgtaagaggg	cctggtttct	tttagtgggg	caatttgcat	gctataaggt
77761	ggtaagtga	tgatttttaa	aagtaagaaa	taacagatac	taggtaatga	ataacagtac
77821	aaaactgata	gaactatgaa	ggtattgcag	cccaagaccc	cattagttag	ggggagatag
77881	tctcacaact	taacaggccc	tgaagggaat	aatgcatctt	aatatcata	cacaccatca
77941	caYggatg	agccttagtc	cccaagaaaa	ccttaaaact	tatttagcaa	acttatttta
78001	gaggtatgtg	taaaacattt	tttaaagttt	catatatgga	agccttagct	gtctggaaaa
78061	ttagtgtaaa	tagaatgctt	catttgctca	ggtttatggt	gcagaaaaa	atcatttagg
78121	gagcagctta	aaatataggt	gtctggacca	ggcaatggtg	gctcacgcct	gtaatcccag
78181	cactttggga	ggccaaggca	ggaggatcac	ctgaggtcag	gagttcgaga	ccagcctgga
78241	caacacagg	agaccccatc	tttcaaaaa	ataaacaaaa	ttagctaggt	ggcacactcc
78301	tgtggtccta	gctacttggg	aagctgaggt	aggaggatcg	cttgaggatg	ggggttgagg
78361	cttaggtgag	gtgtgatggt	accaccacac	tccagcctgg	gcgatagagc	aacctgtctt
78421	caaaaagaaa	aaaaaaaaaa	aagtatggat	ttctgggtcc	cactcYagac	aacttgaatc
78481	tgggtgtagg	gcttcgaatt	tgtattttag	taagccttcc	aatgatattt	gaagtaggta
78541	gtcctaaaaa	ctaacctttg	aggaaacatt	ccagactgac	tgatagaaa	agagagagtg
78601	aataggggaa	gaggagggaa	aatgctttaa	aagagtgtat	cttggttggg	cgcagtggtg
78661	ctgcttgtga	atcctggcac	tttgggaggg	tgaggcgggt	ggatcacttg	aggcaggag
78721	tttgagacca	gcctggccaa	catggagaaa	tctctactaa	aaatacaaaa	aaaaaaaaaa
78781	aaaaaagaaa	gtaaaaaaa	gagtgtatct	tgcagcatat	ttctgaggat	taagccatgt
78841	tttaaaaaatt	tagtgattat	Yggccggggc	cRgtggctca	cgctgtaat	cccagcactt
78901	tgggagggcg	agggcgggcg	atcacgaggt	caggagatcg	agaccatcct	gggcaacagg
78961	gtgaaccccc	gtctctacta	aaaatacaaa	aaaaaaatta	gccgggcgtg	gtggcgggcg
79021	cctgtagtcc	cagctacttg	ggaggctgag	gcaggagaat	ggcgtgaacc	cggggggcgg
79081	agcttgagct	gagccgagat	cgccgacttg	cactccagcc	tgggggacag	agcgagactc
79141	cgtctcaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaatta	gtgattatcg
79201	gccaggccca	gtggctcacg	cctgtaattc	cagcattttc	ggaggctcag	gcaggtggat
79261	cacttgaggt	cagttagttc	agccagcgtc	ggccaacatg	gcaaaaccct	gtctctacta
79321	aaaatacaag	aattagctgg	gtgtggtggc	gtgtgcctgt	aatctgagct	acttgggagg
79381	ttgaggcagg	agaattgctt	gaaccagga	agcagaggtt	gcagttagcc	gagatcgcg
79441	cactgcccct	cagcctgggt	gacagagtga	tactctgtct	caaaggaaaa	aaaaaattgg
79501	tcattctctt	tttctgattt	aaaaggattt	tgttctctca	tggaaaaact	taaaaatagg
79561	agcaagcaaa	aaaaataaaa	aaaatttgta	atcttaccac	tcggagacaa	tgttctattt
79621	aagttggtgt	atattcttcc	agacttctaa	aagggaattg	aattgctcct	ttttagtcag
79681	acaccatcag	ggacttgctt	ggggcctctc	atctagcaag	tgggagagtc	cacactggaa
79741	cccagctgct	gaatcctgaa	cttcattgtg	ctccattgta	ttttttattt	ttattttaaa
79801	aaattaatta	tttatttttg	agacaaggct	ctgttKccca	ggctgtgtgg	tgcagtggca
79861	tgatcacagc	tcatggtagc	ctccatctcc	caggctcaag	cgatcctccc	acctggcctt
79921	cccaaagtgc	tggtattaca	ggtatgagcc	accgcccccc	attgtgcttt	atcgtaacag
79981	gaacagaggg	gctcacgccc	actcctagca	ttactttaca	cacatgtcta	ctgttaaaag
80041	acaaaactac	aacaaattta	gtttaaagat	cttaattggt	ttttatttgc	aattatagaa
80101	tcagccaaca	cctcattcta	gaaaatagaa	tgagtaatga	gtgttctgat	gagctgagca
80161	ggagagggtg	gcttttatag	tagaagaggg	ctgaagaaa	caaaaatggg	ccgggcgcga
80221	tggctcagac	ctgtaatacc	agcactctgg	gagtctaggg	cggacagatc	acttgagttc
80281	aggagtctga	gaccaacctg	accaacatgg	tgaaccctg	tctctaccaa	aaatacaaaa
80341	attagctggg	tgtgatgggt	cacacctgta	ttcccaagta	ctcgggaggg	tgaggcagga
80401	gaatcacctg	aacctgggag	acagaagttg	cagttagctg	agattgtgcc	agtgactcc
80461	aggctgggtg	atagagttag	acctgtcttc	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa
80521	gaaagcaaaa	acagagaaca	aaatggagat	tggttgtttt	caagttactt	tccttaagaa
80581	gttaaaacag	aggtgattac	tttattatgc	tgactccggt	tgactggaat	cttctgtttt
80641	gttttgtttt	aaaagagaga	tggagtcttg	ctttttggcc	aggttggtct	ggaacttctg
80701	acctcaagca	agtcccccac	cttggccttc	caaagtgtcg	ggattacaga	agtgaaacc
80761	catgcccgcc	tggaccctcc	tattttgttt	tggtttttta	ggaaaactgg	cctgttttag
80821	tttactgtgt	tggcacttag	cacatgtgac	tccattcttg	tttggcctgg	tctgatgggg
80881	cccagtgcag	gaggctgggt	caaataYgg	cctccatta	actttgcttc	acaccgcaca
80941	tcacttggtt	aattaactgc	acaattacaa	caaaaaataa	tgtgcctact	actgaactat
81001	gggttcttca	gtagggcagg	aacatcatct	tatttatatc	agtatccttg	gaacacagcc
81061	tgggtcttaa	ctgtggctca	atacctgtaa	cacagtaaat	acacaataaa	tgatccttct
81121	gggcaggggg	caatggagtg	aaaacaacta	ttctccattg	cctctcacag	aactgaattg
81181	cagccatcct	agacttcaaa	actgatctag	ttttctcccc	aaaaaacagt	ttgcagcact
81241	ccaatgtcca	agaaatagaa	tatcctaaac	gaattagggc	aggaacaaaa	aaccaaatat

81301	cacatattct	catttgaaag	tggaaagctaa	acatcagata	ttcatggaca	ttaggatgac
81361	aatggtagat	gcggggagact	actagagtcg	gggaggagag	agggggcaag	ggttgaaaaa
81421	ctgttgggta	cctgggtgac	gggatcattc	ataccccata	cctcagcatc	acgcagtata
81481	cccaggtaac	aaacctgcat	atgtacccSc	taaaaactaaa	atccagattg	aaaaagaagg
81541	aaaaaaaaaga	agaaacagaa	tatcttctctg	aaagccttat	aggggaacat	gcagaagaaa
81601	gcaaatcctc	ttagcatctt	accccttac	ctaagcaaca	gggcccctga	cataaatggg
81661	atgaccgcgc	tccagggttg	gatttgaaag	ccagcgcac	gtgctgcaat	aaaggctcac
81721	caggttttaga	aacaaaagcc	tggctgtctc	cttttctctc	ttagcctcac	aaacagctac
81781	tacttggtg	ctgagtggtt	gaaacaaaag	catacaggag	gaataaaggt	ggatttaact
81841	gacagtcttt	tggaatgtaa	ataccaatca	aatgaagaa	atatgttttt	cagcttcaat
81901	acaatctact	ttgggtttatg	agagtgaaaa	gtctgatgaa	atacatagtt	ataaaaaatgt
81961	gaaaactcct	aaggaagtaa	atctgtacag	gagaattggg	gagctgaaat	atttccaggt
82021	actgttttat	agtcttcttt	ttctttgttt	ttcttctttt	ttctcccagg	aaacatttta
82081	ttggatgtgt	tcatttatgt	atgcacgtat	gtgttcttcc	aaatgggtgt	tttctctggaa
82141	ccaacttcat	gtgtgggtag	cactgtgcta	gatactagct	gcataataaa	tgaatgtgaa
82201	agttgttgac	atataggact	gatttttact	attcattgac	cccttcgaac	atttatttgg
82261	ctgctatgat	tctatgctta	gaaagataca	tatatctaag	caaatacata	caatttttag
82321	gagttcatga	tccctagaaa	gcctcccagg	ggtcaacaaa	ccagaactgg	ctgggagcag
82381	tggctcatgc	ctataatgag	tactcctgag	gtgggaggat	cactggagcc	taggagttca
82441	agaccagcct	agacaacatt	gggagattcc	atttctacaa	aaagaaaatt	atctgggtgt
82501	ggtgtgctg	gcctgtagt	ccagttactt	gggaggctga	ggctggagga	cctctgagtc
82561	ctggaaggtc	gaggctgcag	tgagccatga	ttatgtcact	gcactccagc	cagggaacaa
82621	gtgagactct	gtttcaaaaa	agaaccccca	aaacaaacaa	aaacccaaca	gaatcctgct
82681	ctaagaacct	atagttaaaa	aagcagcagt	gtaggctggg	cacggtggct	catgcctgta
82741	atctcagcac	tttgggaggc	cgaggtgggc	agatcacttg	aggccaggag	ttcaagacca
82801	gcttggcgaa	acactgtctc	tactaaaaaa	tacaaaaaatt	agctgagtgt	ggtgatgcac
82861	gccggtaatc	ctagctactc	tggaggtgta	ggcacaagaa	tcaattgaac	tcaggaggtg
82921	gaggttgtag	tgaactgaga	ttgcgccact	gcattccagc	actccagcct	gggtaatgag
82981	agactccgtc	ttaaaaacaaa	acaaaaacaaa	acaaaaacaaa	acaaaaaccaa	gcctgtaatc
83041	atggctgagt	aataaaaaag	gttttggctg	ggcgcggtgg	ctcatgcctg	taatcccagc
83101	actttgggag	cttgaggcag	cagcagatca	cgtggtcagg	agtttgagac	cagcctgacc
83161	aacatggtga	aaccccgctc	ctactaaaaa	tacaaaaaatt	agctggatgt	ggtggtgggt
83221	gcctgtaagc	ccagctgctc	aggaggtgta	ggcaggagaa	ttgcttgaat	ctgggaggcg
83281	gaggctgcag	tgagccaaga	ttgtgccact	gcactccagc	ctgggtgaca	gagcgagact
83341	ccatctcaaa	aaaaaaaaaaaa	aaaaaaaaaaaa	aaaaaagctt	tgaccaatgc	aagatcagtg
83401	tgtgtatagag	tggttagaag	aaagtccacag	agttcggccg	ggcatggtgg	ctcatgccc
83461	taatcccagc	actttgggag	gccgagatgg	gtggatcaca	aggtcaggag	atcgagacca
83521	tcctggctaa	cacggtgaaa	ccccgtctct	actaaaaata	caaaaaaaaaa	ttagcYgggc
83581	ggggtggtgg	gtgcctgtag	tcccagctac	tcaggaggct	gaggcaggag	aatggcgtga
83641	acccgggagg	tggagcttgc	aStgagccga	gatcgcgcca	ctgcgctcca	gcctgggtga
83701	cagagcaaga	ctcgtctca	aaaagaaaaa	aaaaaagaaa	gtcacagagt	tgaccttaaa
83761	ggctgggcaa	gatttggagt	gYgcagtaga	cagaaggagg	caggggtagg	gccagcaaa
83821	gagcaaggag	gcagacagca	tggaaatagg	gaggggcccag	ctggaaggag	gtcatgttca
83881	tggcaagtct	cggttcggac	ttcacctgat	agattgacca	gggtttctcc	acagtgtatc
83941	tgttgacatt	tggaaactgga	cagttctttg	ttgtgaaggc	cagttctgtg	cattgtgaaga
84001	tgccaacaaa	cacacttgac	ctctgcccac	tagatgccag	tagcatctcc	caagtcatga
84061	caatcacaaa	tgtctccaga	cattgccaaa	gatccccctg	gggacaaaaga	tcaccctatt
84121	tgagtatgac	tgtcacagca	tatttgagct	tggcaatggt	ttccaaagtt	tagtcattca
84181	ggtatcacct	tcacgaat	tgccagatct	tgggtgctcc	ttcttggcta	tctcctgtat
84241	ttattttatt	atttagggac	agaggctgag	gcaggagaat	cKcttgagcc	caggaggcag
84301	aggttgagct	gagccgagat	caggccactg	cactttcagc	ctgggtgaca	gaatgagact
84361	ccatctccaa	aaaaaaaaaaaa	aaaaaaaaaaaa	agtaaggatt	tacttctcct	cagatgttgg
84421	caaaaaaggt	ggaggccaaa	gacRattggc	ctcttccctc	tacagagagg	ccaactcatt
84481	ccctgtgcca	tcatcatggt	agtagctctc	ggggcagatt	gcaagacccc	tcctctttca
84541	ggattgaatt	cccccttggc	atgtgtcaca	tcttcatttg	cctttctttt	ctgatgagag
84601	ggtcatcacc	acttctctgat	ctcttcttct	ggcaacagaa	gggtgatgtg	agattgaatt
84661	tctcaaaaaa	actaaaaaag	atgccaaagt	ccccaggatc	cgtcccacgt	gggccaagat
84721	cttcttttgg	tgcagagggt	cagtgcagac	aaaggcactt	ggatcgcttg	cgatgggtcc
84781	tctctctcta	attatagaaa	gtatttagct	attattccct	aattgctcac	tgcagttggg
84841	tgttagcatt	tttttttttt	tctgacaaaa	ggctgggttc	tgatgagttg	ttcttttctt
84901	tcgcgcttgg	tctactgccca	taccagtgcc	tggctcttcc	actggaacac	tgcaggaggt
84961	ggtatggtat	ttctggggaa	gccactttac	ctcatccaat	tgctttcatg	gcttcccaggt
85021	tgccataatag	gaagggaatt	aacattttact	gggcatctat	ttagtatgtg	ccaggcactg
85081	tgctaagtgc	tcacgtaggt	tgtcttattc	aatccttgca	gcaaacttSa	gaggcagata

85141	taattattcc	Raatttatag	atgaagaatc	taagggtgaag	gtagtctcag	tcagtatgtg
85201	gctgagtcag	aattcaaacc	caggctcttg	tgtgtgagtc	cttttttttt	ttttttttga
85261	gacatgggtc	cactctgtca	ctcaggctgg	agtgcMgcgg	tgcaatcttg	gctcactgca
85321	gcttcaacct	cctgggctca	aacaattctc	ctgcctcagc	ctcccaagta	gctgggacta
85381	cagggtgtgtg	ccaccacaca	tggttaatct	tgttttattt	tttaatgaga	tgggctctca
85441	ccatggtgcc	caggctgggc	ttgaacttct	gggctcaagc	aattctcctg	ctggagctcag
85501	tctgggcaca	gtggctgagg	cttgtaatcc	cttcccaaaa	tgctgggatt	gtgagctact
85561	gtgcccagac	tgactccatg	ttttcctctc	tcaaaagttt	ttctagactt	ctggctgata
85621	aaacattggg	gttgagagact	cccgcctctc	tgccctaata	gtgcttattt	ctaggtcctc
85681	tgctagtggg	agggtctggg	ggggtatgct	ataggacgct	ttaaaggagct	ctaggtttcc
85741	tgttttgggt	catgacacat	gaaggaattt	ccagtgtcat	ggtagcaagg	acgaggagga
85801	aagccatcSc	tcctggggct	gtagcagtaa	ggatcctatg	ggaaacagat	ggcacattca
85861	ggggtttaca	tgacagaacag	aataattcaat	gaatggacta	cttagcagag	ttatgggcag
85921	ttaaaggaac	caacaagcaa	tggggaagct	accgggtact	ggcagcagca	ggaagccatc
85981	accaccccta	ggtttaaagg	gacatgggtg	gagaatgggtg	ttcctgaagc	ccattgacag
86041	ctgggggtat	ggcagaagca	tccagacagg	agctgtgaca	ggagggaaac	ccagccactg
86101	cccgaacacc	agcaggaagg	gcagcagcag	ctgggtgtgtg	agtgtgtata	tatacacaca
86161	tccaggtctc	tccctcctct	gccttcccat	ctgctcatgc	taccactggc	caaaccagc
86221	tagaagacag	agagtagggg	agcctagaga	tgacagctct	gcaggtcagc	ctcctgaggc
86281	acagggcagg	aaggaagggt	ggagagagca	cctgggaaga	tacagagagt	gttcagcaca
86341	ggggctgact	caactgtgcc	atccagacga	gtgtcttttt	ttttttNgag	atagagtctc
86401	tctctgttgc	ccaagctgga	gtgcggtgat	gcaatctcag	ctcactgaaa	tctttgcctc
86461	ccgggttcaa	gcgattcttc	tccctcagcc	tccYgaatag	ctgggactac	ataagtgcac
86521	caccacaccc	agctaatttt	tgtattttta	gtagagacgg	ggttttgcta	tgttgcccag
86581	gctggtctca	aactcctggc	ctcaagtgat	ccaccctcct	tggcctccca	aagtgcataag
86641	attacaaatg	tgagtcacca	cacctggccc	agtgtcatat	ttttgggtgag	tgagagcata
86701	ctctatagga	agtgggacat	tgtttacagg	ctcaaagaat	catggaatgt	gagagtggga
86761	agtgccttta	gagctctacc	tgtaaaaatt	ctttacttta	taaagaggat	cctgaggccc
86821	aggaagggga	aaggactggg	ccaaggtcag	ttgctaattc	gtggcatggg	caggactgga
86881	agccagcttg	cctaacttcc	agtccatact	gcttcttagt	tttgagctgt	agtcaattaa
86941	tggttggtaa	taaaggatta	ctactgagcc	acttgaccta	tttctctgta	agacctcagg
87001	gtatagttta	ctttggcccc	agacatttgc	ctcaaataat	tctttctcta	aataatctca
87061	tgatcagagt	aggctgcccc	tttcagagtg	tgggctctgg	agtcaggctc	atggggctga
87121	atcccagccc	tatccttaca	gaatgggtct	ggacaagtta	tttaaccctc	caccttattt
87181	tgtcactg	taaaatgagg	ataataatga	tgtctatctc	atatctcatg	aggattatat
87241	aagaagatat	atttaaagca	cttagactag	catctgcca	taatgagtgc	atgataaata
87301	acagtataca	tgtgggggtg	gtgtgtatata	atataatga	tgtaaacaaa	tatgtgtgtg
87361	tgtgtatgta	tatgcatgta	tgtgtatata	tatgcatgta	tgtgtgtgtg	tgtgtgtgtg
87421	tatatatata	tattctaccc	cccaaattag	tttagaattc	tggttaagaa	ttcacatggg
87481	tgggtgggtg	gtgggtggctc	atgtctgtaa	tcccagcact	ttgggaggct	aagggtgagca
87541	gatcacttga	gtgcaggagt	ttgagactag	cctggccaac	atgggtgaaat	cctgtctcca
87601	ctaaaagtac	aaaaattagc	cagggtgtgat	ggatgtgtgc	tgtagtccca	gctacttggg
87661	aggctgaggc	aggagaaacg	cttgaacctg	ggaagtggag	gttgagctga	gccgacattg
87721	cacactgcac	tccagcctgg	atgacaaagc	gcgactccat	ctcaaaaaaa	aaaaaaaaaa
87781	gaattcatgt	ggttggctgg	gcaaggtggc	ttatgcctgt	aatcccagca	cttagggagg
87841	ccaaggcagg	tggattgctt	gagctcggga	gtttgagacc	cacctgggaa	acttgatgaa
87901	ataccatctc	tatcaaaaat	acaaaaaaaa	aaaaattagc	tgggcattgt	ggtgtgtgtg
87961	tgtggtccta	gctacttggg	aggctgaggg	gggaggatct	cttgagcccg	ggaggtggag
88021	gttgacagtga	accaagattg	tgccattgca	ctactccagc	ctgggcgaca	gagaccacat
88081	ctaaaaaaaa	aaaaaaaaaa	aagaattcac	gtggctgtga	ttgaggtcag	gcatgggtgg
88141	gtgtgcctat	agtcccagct	acttgggagg	ctgaggcagg	agaaacgctt	gaacctggga
88201	ggtggagggt	gcagtggacc	gacatcacac	actgcactcc	agcctggatg	acaaagcgag
88261	actccatctc	atgcctgcac	tttgggaggc	tgagttagga	ggatggcttg	agcacagggg
88321	ttcgagatca	gtctaggcaa	catagggaga	cactgtcttt	gcaaaaaaat	acaaaaatta
88381	tctgggcatg	atgcaccacg	cctgtagtcc	cagctacttg	ggaggctgag	gtgggaggat
88441	tacttgagcc	tgggaggtca	aggctgtagt	gagtgcgcca	tgattgcgcc	actgcttctc
88501	agcctgggca	atacaacaag	accctgtctc	aaaaaaccaa	ccaaccaacc	atacttttta
88561	cccttattga	aatatatcag	ttgttctttt	aaacattttt	ttcttttaat	ttttttttta
88621	gtagagacaa	ggtctcacct	atgttgccca	ggctggtctc	aaactcctaa	gctcaagcta
88681	accttctacc	tctgcctccc	aaagtgttgg	gattacaggg	atgagccatg	gctcctggcc
88741	ttcttagctg	tttttcacat	ggaatttttc	ctttccctat	aacagataag	aactctgaag
88801	ggcttaggtg	atgtgccttt	tctgggattc	ctcagcagtc	agaggcagat	aaatccttgt
88861	tctcaggctg	ggcactgtgg	ctcacacctg	taatcccagc	actttgagag	gctgaggcgg
88921	gcggatcacc	tgaggctcag	agtttgagac	cagcctggct	aacatggtga	aaccagtcct

88981	cgactaaaat	acaaaaatta	gccaggcatg	gtgggtgcatg	cctgtaatcc	cagttactcg
89041	ggaggctgag	gcagggggaat	tacttgaact	tgggaggcgg	aagttgcagt	gagccgagtt
89101	tgtgccgatg	cactccagcc	tgggcaacag	agtgagactc	catctccaaa	aaaaaaaaag
89161	aaaatacaag	tctgactcct	cacatcataa	gctggctgtg	gtggctcatg	tctgtaatcc
89221	cagcactttg	ggagggctaaa	gcggatggat	catttgaggc	caggagtttg	agaccagcct
89281	ggccaacatg	gcgaaaccct	gtctctacta	aaaatacaaa	aaatcaggcg	ggcgtgggtg
89341	caggtgtctg	tgatcccagc	tacttgagag	gctgaggcag	gagaatcgct	tgaaccgggg
89401	aggcagaggt	tgacgtgagc	cgagatgggt	ccgttgactc	ccagcctggg	tgacagaatg
89461	agactcggtc	tcaaaaacaa	caataacaac	aacaacaaca	acaacaagc	aagtttgagt
89521	cttcacatta	agcctgtacc	actggtgtaa	catgggaaga	catgaagaag	aatgatctg
89581	agctttaatc	atttatattc	ggacataagg	tccaccaaga	taaagaacat	ctgtaccaaa
89641	ctgccccctc	cccacaaaat	tcaagtattg	ataagtattg	ttaaataaat	tttataaata
89701	aaacagtgtg	aggaggtaga	gtgtgtctca	gggaatgtgg	gggtgagtgg	ggagataggg
89761	gtaattgggg	atgacctctc	tggagRgtaa	cattagcaga	gaacaaattg	ccaggatgga
89821	gcagaccatt	gaagatatgg	gggaatagca	agtgtaacgg	ctgccagat	agcagccagc
89881	ttggcagatt	gagWacagaa	aggaagctga	tgcggttggg	gtggagttag	ccagggaga
89941	gaagggtggg	aaatgaggct	ggagagaggg	agtggagcca	gcctctgggg	agccaggagg
90001	gtcaggctga	aggggtgggt	tgccccctca	cacctgtggg	tgtttctcgt	aagggtggaac
90061	gagagacttg	gaaaagaaaa	agacacagag	acaaagtata	gagaaagaaa	taaggggacc
90121	cKggggacca	gcgttcagca	tatggaggat	cccgccagcc	tctgagtttt	cttagtattt
90181	attgatcatt	cgtgggtggt	tctccgagag	ggggatgtgt	cagggtcaca	agacaatagt
90241	ggggagaggg	tcagcagaca	aacacgtgaa	caaaggctct	tgcatcatag	acaaggtaaa
90301	ggattaagtg	ctgtgctttt	agatatgcat	acacataaac	atctcaatgc	tttacaaagc
90361	agtattgctg	cccgcattgc	ccacctccag	ccctaaggcg	cttttcccct	atctcagtag
90421	atggaacgta	caatcggggt	ttataccgag	acattccatt	gcccagggac	gggcaggaga
90481	cagatgcctt	cctcttgtct	caactgcaag	aggcatgcct	tcctcttata	ctaactctcc
90541	tcagcacaga	cccttttacag	gagttgggct	gggggacggg	cagggtctttc	ccttcccacg
90601	aggccatatt	tcagactatc	acatggggag	aaaacttgga	caatacctgg	ctttcctagg
90661	cagaggctccc	tgccggccttc	cgaggttttt	gtgtcccctgg	gtacttgaga	ttagggagtg
90721	gtgatgactc	ttaaggagca	tgctgccttc	aagcatctgt	ttaacaaagc	acatcttgca
90781	cgcccttaa	tccatttaac	cctgagtttc	acacagcaca	tgtttcagag	agcaggggtg
90841	taggggtaag	gtcacagaat	ctcaaggcag	aagaattttt	cttagtacat	aacaaaatgg
90901	agtctcccat	gtctatttct	ttctacacag	acacagtaac	aatctgatct	ctcttgcttt
90961	tccccacatc	aggcaaagag	tttggtattt	cttctcagtg	tcattgacact	cctctgaaca
91021	gcaacattca	gcttccaaag	cacttgaccc	attgcacttg	atcctcaca	gcccttttag
91081	gtagatagga	ctgacaatag	tattaaactc	attttatcaa	gacagaaact	aaagctctgc
91141	caaaattaga	aaacctgctg	tgagtcaacc	agttcaaata	agattctttg	actctaaatc
91201	ggggtttttt	acctcagctg	tgcatgaaa	tcctacaggg	agctctacaa	aatacctatg
91261	actcagctccc	attcccagta	gttgaagggg	tgctagtggc	agctatggcc	caggaagact
91321	tttaaagctc	tccagggtgat	tttttttttt	agactgggtt	tttctctgtc	accagggtg
91381	gaatgcgggtg	gtgggtgcaat	catgcctcac	tgtaacccttg	accttctggg	ctcaagttat
91441	cttcccacgt	gagcctcctg	agtagttggg	atttttaggca	tgagccactg	tgccgtggccc
91501	ttattattat	tttatttttt	atttattatt	attttttgag	atggagtctt	gctctgtcac
91561	ccaggctgga	gtgcagtgc	atgatctcgg	ctcactgcaa	cctctgcctc	ctaggttcaa
91621	gcgagtatcc	tgcccttagcc	tctagagcag	ctgggactac	aggcgtgtga	caccacgccc
91681	agctaagttt	tgtatttttt	gtagagacgg	ggtttcacca	tgttggccag	gctgggctgg
91741	aattcctgac	ctcaagtgat	ctgcctgcct	tgccctccct	aagtgtcggg	attacagggtg
91801	tgagccactg	cacctggcct	aatgctccct	tttttctgtt	cctaagccac	tttggttcata
91861	cctcaatcct	aacacttctc	acattatttg	gaagttaatca	atcttcttat	cttaatgcct
91921	atacattgca	aacttttagat	ggtgggggct	gtgttattta	actctatact	actgctacct
91981	ctgtgctgta	attatctttt	tttttctttt	ttagtggaccg	ggtttagctc	tctctctggt
92041	gcccagctgg	agtgcagtgg	caagacYcat	caaagctcac	tgacgccttg	aactcctggg
92101	ctcaagatcc	tccctccctg	gctccagag	tagctgggac	tacaggcatg	cgccaccatg
92161	cccagctaat	ttttaatttt	tctgtaRtct	catagggctc	ggctatgttg	ctcaggctgg
92221	tctctaactc	ccagcttcaa	gcgattcttc	tgccctggcc	tcccaaaactg	togttggttac
92281	agggataagc	cactgtgctc	agtcattgct	gYcttttcaat	aaagacaagt	gcttaaatgt
92341	aatgtagttt	aaatttatca	aagttttatt	tacagtttgt	gattatggca	cttaataaac
92401	cctaccctga	gatcataagg	atgttttctt	ttcttgtaaa	agttataatt	ttggggaacg
92461	atagtaagaa	gaaaaaaaaa	gttgtaattt	taaaatgctt	taatttcccc	tgtaagtctt
92521	taacctacct	tgaatttagtt	tttgcatagg	ggatccaaca	ggtaactggt	tgttccagcc
92581	catgtaactg	atactctatc	ttccctcact	gaactataat	tccagccata	cctcatttca
92641	tagatgcctg	catgtttggg	ctctgtagggt	ggttccattt	ccccctctgc	ctttccctat
92701	tccaatatca	catgggtctta	atttgatatg	ctttataata	catgttgaca	tggaaggggtg
92761	cgtccttgaa	acacattttt	cttcttcaag	aatgttttgg	ctatgctttg	ccttaaagtg

92821	gtgKtttttt	tttgtttttt	gtattttttt	ttttttttga	gacaaagtct	cgctctgtca
92881	cccaggctgg	aattcagtg	tatgatcttg	gctcactgca	acatttgtct	cccgggttca
92941	aggattctct	tgccacagcc	ttctgagtag	ctgggattac	agggtgtcac	caccatgcct
93001	ggctaatttt	tgtatttttag	tagagatggg	gtttttaccat	gttggccagg	ctggtgatga
93061	actcctgacc	tcaactgac	cacctgcctt	agcctcccaa	cgtgctggga	ttacaggcac
93121	gagccactgc	gctgcccttt	gttttttcat	acacactttt	gattagtttg	tgaaggcca
93181	gaaaaaatta	tggttgaaat	ttgattggaa	ctgcattgga	ctgatagatt	aattcccca
93241	tgaaaaaaat	agctcaagag	atgtttatag	aataaacaaa	cgtctgaggc	ctggcggtgt
93301	ggctcacacc	tgtaatccca	gcactttggg	aggccaaggc	cagaggatgg	cttagcacag
93361	aaattcgaga	ccagcctggg	caatatagtg	agaccacacc	ccccagcttt	tcttgagaaa
93421	aaatttaaaa	aatttagcca	ggcgtggcag	tatgtgcttg	tagtcccagc	tactccacag
93481	actgaggcag	gagggtacct	tgaggacagg	aagttgaggc	tacagtgatg	gagcctctac
93541	actccagctt	gggcaacaca	gccagaccct	gtctcaaaaa	aaaaaagaga	aatgtttttg
93601	atacatgtct	tcaaagtgca	tcttataagt	caacatttga	taatggagat	cagtgaatga
93661	atgaataagg	gccaaagata	actaagatta	gctagcatat	tatatatttt	tatttattta
93721	ttttgagacg	agYcttgctc	tgttgccag	gctggagtgc	agtggcccaa	tctcagctca
93781	ctgcaacctc	cacctcctgg	gttcaagcaa	ttctcccacc	tcagcctccc	gagtagttgg
93841	gataacaggt	gcccggccact	atgctgggct	aatttttctt	tttctttttt	ttcttttttg
93901	agacagagtc	tcaactctgtt	gcctaggctg	gagtgcagtg	gcgtgatctc	agctcactgc
93961	aacctccgcc	tcccgggttc	aagcgattct	cctgcctcag	cttctctgag	agctgaaatt
94021	ccagactatt	accaccacac	caggctaatt	tttgatattt	tagcagagac	agatttcaact
94081	atgttggcca	ggctgggtctc	aaattcccgg	cctcaagtgg	tccgcccacc	ttggcctccc
94141	aagggtgctg	gattacagga	atgagccacc	gcaccgggct	taatttttct	attttttagta
94201	gaaatgggg	ttcactatgt	cggcctggct	ggctcccaac	ccctagcctc	agggtatctg
94261	ccgcctcag	cctcccaaag	tgctgggatt	acaggtgtga	gctactgctg	cctgcccagc
94321	cagctattatt	attattattt	ttgagacgga	gtttcgcgtg	tggtgcctag	gctgctagtc
94381	aatggcggtga	cctcggtctca	ccgcaacttc	tgcttcccgg	gtttctagt	cctcagcctc
94441	ctgagtagct	gggattacag	gcattgcacca	ctaggcctgg	ctaatttttg	ctattattag
94501	tagggacggg	gtttctccat	gttggtcagg	ttggtctcaa	actcccagacc	tcagggtgatc
94561	tgctagcctt	ggcctgctgg	gattacaggc	gtgagccacc	gccccgccag	catacatcaa
94621	ttgtcagtaa	atttcttctt	cgctactag	ttccaatata	ttattttaa	tattatggcc
94681	ggtgaggttg	ctcaggcttg	taatcccagc	actttgggaa	gtcaaggcgg	gtggattgcc
94741	tgagttcagg	agttcaagac	cagcctgggc	aaaatggcga	aaccccgtca	ctcctaaaaa
94801	tacaaaaaat	tagcagggcg	tggtagcaca	ctcctgtaat	cccagctact	caggaggctg
94861	aggcaggaga	atcacttgaa	cccgggaggc	ggagggtgca	gtgaggcgag	atggggcact
94921	gcactccagc	ttgggcgaca	gagtgcagct	ctgtctcaaa	atgaatgaat	gaatgaatga
94981	atgaatgaat	aaaatctacg	aggtcgggcg	cagtggctca	ctcttgtaat	cccagcactt
95041	tgggcgggcg	gatcatttga	ggtcaggagt	ttgagaccgg	cttgggccaac	atggtgaaac
95101	cccattctcta	aaaaaaatc	aaaaattagc	tgtactgctt	gggcgacaga	gcgagactca
95161	gtctcaaaaa	aaaaaaatct	acaaatccac	ttcatctttc	attaaacacg	aacctgggac
95221	gtagtttagaa	tgctttgaa	aaaatatttc	agtaatatcc	caagtccgtc	gccctggcaa
95281	agcaacaagg	ccacctgacg	gagtaagccc	gagacccctg	tgcttccaca	gtactgcccg
95341	cgggatcgcg	gcgtggggaa	ccaggaaacta	caagtcccgg	caggccgcgc	gcgcctcggc
95401	ttcactaagc	gcagtcagg	ggctgtcccc	gcacggttcc	cagctaattc	cccgtaccg
95461	ggttgcggcc	ggaagccggg	gcgcgggct	ctgcttccct	cggggtgagt	aggggcagct
95521	cgccggaggc	cgccgcagag	ctcgggcgca	ccagcctggc	agggccgcac	cgccctgga
95581	gcgcgaggga	accggcatgg	acacgcgggg	agtcggcctc	tgctccggtt	gtgcagccgc
95641	cctgggcgcc	gggagggaca	gccccatgac	cttgtcacct	gggcccggcg	cgccggccacg
95701	tggtcctggg	gactcgcatt	gccccgcccc	cctgcagtga	cttctccagt	ttggcacacc
95761	agtgggcttt	catgtgacct	gtccaattca	gtggcacagt	cttgaggca	cccagtgagg
95821	cagagctgga	ggagggccgg	aggtggctca	cccagacgcc	ttcacccctt	ggtggccctg
95881	tgagttccga	ggcagaagca	gctggagctg	gagcgtgggtc	tgggcaagag	cagggggcga
95941	tgcccatatc	ctggccctcc	acccggtctt	gcgaacgtgc	cccagaccgc	tgcttatctc
96001	cagcatcttg	cagggccctg	ggcttccaga	tcacacagta	gatatgtgct	aacagcagcc
96061	agtgtgtgtg	gagagcaacc	tttagataag	acacctagaa	cagccctctc	ccagggttaa
96121	gtggcattta	ctttctacta	acaagaaagt	cattcaacaa	gtattttatg	tttactcca
96181	tgatacttgg	cacttagtag	gtgctcaaga	aataatttga	gagttgaagt	gaattagact
96241	ttatctgtgg	cttcatatgt	tttatttaat	aattttcttg	gggctggatg	tggtggctga
96301	tgctgtgaat	cccagcactt	tgggaggcca	agatgggagg	attgcttgag	ggcaggagtt
96361	caacaccagg	ctggggcaaca	tagtgagact	cccgtctgta	ttaaaagaaa	aaaaggaaaa
96421	taaaaaaaat	aattttctta	ctacagtgg	tctcaaaact	tagcatgcac	cagagtcacc
96481	tgagagagcta	gtaacaacac	acagtactgg	accccatccc	caaagtttct	gattctgtaa
96541	gtctgggtatg	agSaccaaga	attggcgttt	ctaacaatga	cccaagtgat	gcagttgctg
96601	ttggtctggg	gaccatacgt	tgagtaccac	tgctttatag	gaaaaaaaat	cctgtgaact

```

96661   taaagcacta tctcttcaca cctgggttacg ggtctttctg tctttatttg ccagttttac
96721   ttttttggtg attgtaaata tgtgcatatt actttgtaca ctgttaactc cctccatatt
96781   aatttatata cagattttta ttagtcttat acttgtcatt tttcctggtt acatcatatt
96841   ctaccttatt gggtcattct ttgcttggcc atttctcaac tggcctcagg gatgtgcat
96901   ttgatctgta ctagaagagt gtgcatattc ttactttctg Mtatgcagtc attttcttgt
96961   cagacaaagg acctgtaaga atattagttg gggccgggtg cggtggctca cacctgtaat
97021   cccagcactt tgggaggctg aggtgggcga atcacttgag gtcaggcggt caatgccagc
97081   ctggccaaca tggtgaaacc ccatctctac caaaaataca aaaaaattac ctgggcgtag
97141   tgggtgtgtc

```

[0256] Modifications may be made to the foregoing without departing from the basic aspects of the invention. Although the invention has been described in substantial detail with reference to one or more specific embodiments, those of skill in the art will recognize that changes may be made to the embodiments specifically disclosed in this application, yet these modifications and improvements are within the scope and spirit of the invention, as set forth in the aspects which follow. All publications or patent documents cited in this specification are incorporated herein by reference as if each such publication or document was specifically and individually indicated to be incorporated herein by reference.

[0257] Citation of the above publications or documents is not intended as an admission that any of the foregoing is pertinent prior art, nor does it constitute any admission as to the contents or date of these publications or documents. U.S. patents and other publications referenced herein are hereby incorporated by reference.

What is claimed is:

1. A method for identifying a subject at risk of osteoarthritis, which comprises detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in a nucleic acid sample from a subject, wherein the one or more polymorphic variations are detected in a nucleotide sequence selected from the group consisting of:

- (a) a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
- (b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
- (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
- (d) a fragment of a nucleotide sequence of (a), (b), or (c);

whereby the presence of the polymorphic variation is indicative of the subject being at risk of osteoarthritis.

2. The method of claim 1, which further comprises obtaining the nucleic acid sample from the subject.

3. The method of claim 1, wherein the one or more polymorphic variations are detected within a region spanning chromosome positions 31118000 to 31129000 in human genomic DNA.

4. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 1 selected from the group consisting of 247, 1535, 2386, 6440, 9133, 9143, 9471, 13150, 13717, 14466, 15769, 16870, 18545, 18749, 19123, 20736, 21038, 21046, 21050, 21056, 21706, 23170, 25028, 27871, 28070, 31717, 32019, 32318, 33080, 33101, 34236, 34285, 34818, 35168, 37981, 38113, 38117, 38481, 38615, 38944, 39288, 41385, 42136, 42185, 42353, 42434, 44580, 44675, 45739, 46439, 47457, 47735, 50319, 50708, 51185, 53002, 53064, 53637, 55274, 55825, 55986, 56684, 57653, 57659, 57692, 57775, 61313, 61431, 61699, 62906, 63619, 64664, 68452, 69665, 69681, 70091, 74637, 74760, 76523, 78559, 79549, 79882, 81339, 81681, 81696, 83517, 85431, 86332, 87358, 87725, 89052, 90020, 90231, 90284, 90447, 90601, 90724, 92559, 95176, 95195 and 96822.

5. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 1 selected from the group consisting of 13150, 21046, 23170, 25028, 44580, 62906, 64664 and 83517.

6. The method of claim 1, wherein the one or more polymorphic variations are detected within a region spanning chromosome positions 36914000 to 36931000 in human genomic DNA.

7. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 2 selected from the group consisting of 211, 7217, 7895, 13308, 14279, 17026, 18271, 20417, 21843, 22069, 22145, 22519, 22539, 23236, 23256, 23402, 23499, 23620, 23871, 24136, 25427, 25866, 26541, 26576, 26689, 26720, 27113, 27164, 27186, 28341, 29160, 29844, 30665, 30830, 31061, 31523, 32326, 32346, 32358, 34909, 34975, 35066, 35096, 35375, 36304, 36712, 36770, 37342, 37412, 37884, 38077, 38300, 38301, 41189, 44408, 44493, 44571, 44670, 45219, 45258, 47261, 48473, 48771, 55292, 56479, 56747, 60620, 60688, 61058, 61129, 61577, 61961, 63351, 63926, 65798, 66043, 66044, 66246, 66318, 66547, 71238, 71283, 71492, 72274, 73762, 74209, 75284, 77347, 77589, 78096, 78606, 78862, 79135, 79146, 79456, 79609, 80086, 80119, 80766, 81110, 81269, 81668, 82433, 82559, 83298, 83821, 84121, 84147, 84543, 84554, 84691, 84727, 85678, 86699, 86700, 86792, 86832, 87045, 87140, 87365, 88342, 88498, 88589, 95502, 96968, 97448, 97568 and 98724.

8. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 2 selected from the group consisting of 23236, 32358, 47261, 48771, 55292, 60688, 72274, 74209, 77589, 79135, 79456, 79609, 80119, 80766, 81110, 82433, 84121, 84147, 85678, 86699, 86832, 87140 and 88589.

9. The method of claim 1, wherein the one or more polymorphic variations are detected within a region spanning chromosome positions 76196500 to 76221500 in human genomic DNA.

10. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 3 selected from the group consisting of 218, 1440, 1442, 2611, 4317, 4724, 4788, 5202, 5780, 5974, 6644, 7430, 7938, 8095, 8183, 8312, 8352, 9348, 9378, 9617, 9727, 9834, 9899, 10211, 10377, 10695, 10729, 10730, 11433, 11951, 12697, 12982, 14419, 14501, 14983, 15280, 15475, 15888, 15976, 16307, 16442, 17255, 18948, 19435, 19753, 20021, 20022, 20503, 20590, 21804, 21919, 21990, 22412, 22536, 23432, 23468, 23772, 24325, 24773, 26274, 27440, 28561, 30071, 31764, 33008, 35310, 35460, 37112, 37285, 37747, 38057, 38859, 38860, 39525, 40216, 40281, 41453, 42091, 42513, 42935, 42985, 43003, 43281, 43716, 43866, 44234, 44596, 44871, 45005, 45282, 47178, 47816, 47887, 48134, 48135, 48276, 48400, 48798, 48803, 49146, 49969, 51059, 51064, 53285, 54560, 54748, 54785, 55102, 55644, 55705, 55841, 56623, 56825, 56827, 56892, 59150, 59958, 60231, 60524, 61871, 62226, 63230, 63468, 63787, 65732, 65989, 68832, 69904, 70365, 70886, 73088, 73103, 75934, 75966, 76273, 77943, 78466, 78861, 78872, 79836, 80908, 81509, 83576, 83662, 83782, 84282, 84444, 85129,

85151, 85296, 85809, 86387, 86494, 89786, 89894, 90122, 92067, 92187, 92312, 92824, 93733, 96553 and 96941.

11. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 3 selected from the group consisting of 4788, 8312, 9378, 9727, 9899, 10211, 27440, 40216, 40281, 42091, 43866, 48803, 51059, 55644, 56623, 73103, 78872, 79836, 85129, 92824 and 96941.

12. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in Table A.

13. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in linkage disequilibrium with one or more positions in claim 4, 7, 10 or 12.

14. The method of claim 1, wherein detecting the presence or absence of the one or more polymorphic variations comprises:

hybridizing an oligonucleotide to the nucleic acid sample, wherein the oligonucleotide is complementary to a nucleotide sequence in the nucleic acid and hybridizes to a region adjacent to the polymorphic variation;

extending the oligonucleotide in the presence of one or more nucleotides, yielding extension products; and

detecting the presence or absence of a polymorphic variation in the extension products.

15. The method of claim 1, wherein the subject is a human.

16. The method of claim 15, wherein the subject is a human female.

17. The method of claim 15, wherein the subject is a human male.

18. A method for identifying a polymorphic variation associated with osteoarthritis proximal to an incident polymorphic variation associated with osteoarthritis, which comprises:

identifying a polymorphic variation proximal to the incident polymorphic variation associated with osteoarthritis, wherein the polymorphic variation is detected in a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation;

determining the presence or absence of an association of the proximal polymorphic variant with osteoarthritis.

19. The method of claim 18, wherein the incident polymorphic variation is at one or more positions in claim 4, 7, 10 or 12.

20. The method of claim 18, wherein the proximal polymorphic variation is within a region between about 5 kb 5' of the incident polymorphic variation and about 5 kb 3' of the incident polymorphic variation.

21. The method of claim 18, which further comprises determining whether the proximal polymorphic variation is in linkage disequilibrium with the incident polymorphic variation.

22. The method of claim 18, which further comprises identifying a second polymorphic variation proximal to the identified proximal polymorphic variation associated with osteoarthritis and determining if the second proximal polymorphic variation is associated with osteoarthritis.

23. The method of claim 22, wherein the second proximal polymorphic variant is within a region between about 5 kb 5' of the incident polymorphic variation and about 5 kb 3' of the proximal polymorphic variation associated with osteoarthritis.

24. An isolated nucleic acid comprising a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation; and

(e) a nucleotide sequence complementary to the nucleotide sequences of (a), (b), (c), or (d);

wherein the nucleotide sequence comprises a polymorphic variation associated with osteoarthritis selected from the group consisting of in SEQ ID NO: 1 a guanine at position 13150, a thymine at position 21046, an adenine at position 23170, an adenine at position 25028, a guanine at position 44580, a guanine at position 62906, a cytosine at position 64664 and a cytosine at position 83517; in SEQ ID NO: 2 an adenine at position 23236, a cytosine at position 32358, a guanine at position 47261, a guanine at position 48771, a cytosine at position 55292, an adenine at position 60688, a guanine at position 72274, a guanine at position 74209, a cytosine at position 77589, an adenine at position 79135, a thymine at position 79456, an adenine at position 79609, an adenine at position 80119, a cytosine at position 80766, an adenine at position 81110, a cytosine at position 82433, a cytosine at position 84121, a thymine at position 84147, a cytosine at position 85678, a thymine at position 86699, an adenine at position 86832, a guanine at position 87140 and an adenine at position 88589; in SEQ ID NO: 3 a guanine at position 4788, a thymine at position 8312, a deletion at position 9378, a cytosine at position 9727, a guanine at position 9899, a cytosine at position 10211, a guanine at position 27440, a guanine at position 40216, a cytosine at position 40281, an adenine at position 42091, a guanine at position 43866, an adenine at position 48803, an adenine at position 51059, an adenine at position 55644, a cytosine at position 56623, a cytosine at position 73103, an adenine at position 78872, a guanine at position 79836, a cytosine at position 85129, a guanine at position 92824 and an adenine at position 96941; and an allele associated with osteoporosis in Table A for positions rs552, rs12904, rs2282146, rs734784, rs1042164, rs749670, rs955592, rs1143016, rs755248, rs1055055, rs835409, rs927663, rs8162, rs831038, rs33079, rs1710880, rs4472029, rs799570, rs1282730, rs1518875, rs1568694, rs905042, rs1957723, rs794018, rs707723, rs893861, rs1914903, rs2062232, rs26609, rs1370987, rs1012414, rs435903, rs1248, rs703508, rs226465, rs241448, rs763155, rs1040461, rs462832, rs804194, rs1022646, rs1569112, rs805623, rs1019850, rs1599931, AA, rs279941, rs1062230, rs1859911, rs1477261, rs1191119, rs657780, rs1393890, rs1478714, rs868213, rs690115, rs1465501, rs899173, rs10477, rs926393, rs465271, rs13847 and rs738658.

25. An oligonucleotide comprising a nucleotide sequence complementary to a portion of the nucleotide sequence of (a), (b), (c), or (d) in claim 24, wherein the 3' end of the oligonucleotide is adjacent to a polymorphic variation associated with osteoarthritis.

26. A microarray comprising an isolated nucleic acid of claim 24 linked to a solid support.

27. An isolated polypeptide encoded by the isolated nucleic acid sequence of claim 24.

28. A method for identifying a candidate therapeutic for treating osteoarthritis, which comprises:

- (a) introducing a test molecule to a system which comprises a nucleic acid comprising a nucleotide sequence selected from the group consisting of:
 - (i) a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
 - (ii) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
 - (iii) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
 - (iv) a fragment of a nucleotide sequence of (a), (b), or (c); or

introducing a test molecule to a system which comprises a protein encoded by a nucleotide sequence of (i), (ii), (iii), or (iv); and

- (b) determining the presence or absence of an interaction between the test molecule and the nucleic acid or protein,

whereby the presence of an interaction between the test molecule and the nucleic acid or protein identifies the test molecule as a candidate therapeutic for treating osteoarthritis.

29. The method of claim 28, wherein the system is an animal.

30. The method of claim 28, wherein the system is a cell.

31. The method of claim 28, wherein the nucleotide sequence comprises one or more polymorphic variations associated with osteoarthritis.

32. The method of claim 31, wherein the one or more polymorphic variations associated with osteoarthritis are at one or more positions in claim 4, 7, 10 or 12.

33. A method for treating osteoarthritis in a subject, which comprises contacting one or more cells of a subject in need thereof with a nucleic acid, wherein the nucleic acid comprises a nucleotide sequence selected from the group consisting of:

- (a) a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
- (b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
- (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;

(d) a fragment of a nucleotide sequence of (a), (b), or (c); and
(e) a nucleotide sequence complementary to the nucleotide sequences of (a), (b), (c), or (d);
whereby contacting the one or more cells of the subject with the nucleic acid treats the
osteoarthritis in the subject.

34. The method of claim 33, wherein the nucleic acid is RNA or PNA.

35. The method of claim 34, wherein the nucleic acid is duplex RNA.

36. A method for treating osteoarthritis in a subject, which comprises contacting one or more
cells of a subject in need thereof with a protein, wherein the protein is encoded by a nucleotide sequence
which comprises a polynucleotide sequence selected from the group consisting of:

- (a) a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
 - (b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in
SEQ ID NO: 1-3 or referenced in Table A;
 - (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the
amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
 - (d) a fragment of a nucleotide sequence of (a), (b), or (c);
- whereby contacting the one or more cells of the subject with the protein treats the osteoarthritis
in the subject.

37. A method for treating osteoarthritis in a subject, which comprises:
detecting the presence or absence of one or more polymorphic variations associated with
osteoarthritis in a nucleic acid sample from a subject, wherein the one or more polymorphic variation are
detected in a nucleotide sequence selected from the group consisting of:

- (a) a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
- (b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in
SEQ ID NO: 1-3 or referenced in Table A;
- (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the
amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
- (d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic
variation; and

administering an osteoarthritis treatment to a subject in need thereof based upon the presence or
absence of the one or more polymorphic variations in the nucleic acid sample.

38. The method of claim 37, wherein the one or more polymorphic variations are detected at one or more positions in claim 4, 7, 10 or 12.

39. The method of claim 37, wherein the treatment is selected from the group consisting of administering a corticosteroid, a nonsteroidal anti-inflammatory drug (NSAID), a cyclooxygenase-2 (COX-2) inhibitor, an antibody, a glucocorticoid, hyaluronic acid, chondrotin sulfate, glucosamine or acetaminophen; prescribing a heat/cold regimen or a joint protection regimen; performing joint surgery; prescribing a weight control regimen; and combinations of the foregoing.

40. A method for detecting or preventing osteoarthritis in a subject, which comprises:
detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in a nucleic acid sample from a subject, wherein the polymorphic variation is detected in a nucleotide sequence selected from the group consisting of:
(a) a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation; and
administering an osteoarthritis prevention or detection procedure to a subject in need thereof based upon the presence or absence of the one or more polymorphic variations in the nucleic acid sample.

41. The method of claim 40, wherein the one or more polymorphic variations are detected at one or more positions in claim 4, 7, 10 or 12.

42. The method of claim 40, wherein the osteoarthritis prevention is selected from the group consisting of administering a corticosteroid, a nonsteroidal anti-inflammatory drug (NSAID), a cyclooxygenase-2 (COX-2) inhibitor, an antibody, a glucocorticoid, hyaluronic acid, chondrotin sulfate, glucosamine or acetaminophen; prescribing a heat/cold regimen or a joint protection regimen; performing joint surgery; prescribing a weight control regimen; and combinations of the foregoing.

43. A method of targeting information for preventing or treating osteoarthritis to a subject in need thereof, which comprises:

detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in a nucleic acid sample from a subject, wherein the polymorphic variation is detected in a nucleotide sequence selected from the group consisting of:

- (a) a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
- (b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
- (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A;
- (d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation; and

directing information for preventing or treating osteoarthritis to a subject in need thereof based upon the presence or absence of the one or more polymorphic variations in the nucleic acid sample.

44. The method of claim 43, wherein the one or more polymorphic variations are detected at one or more positions in claim 4, 7, 10 or 12.

45. A composition comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and an antibody that specifically binds to a protein, polypeptide or peptide encoded by a nucleotide sequence identical to or 90% or more identical to a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A.

46. The composition of claim 45, wherein the antibody specifically binds to an epitope comprising an amino acid encoded by rs734784, rs1042164, rs749670, rs955592 and rs1040461.

47. A composition comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and a RNA, DNA, PNA or ribozyme molecule comprising a nucleotide sequence identical to or 90% or more identical to a portion of a nucleotide sequence in SEQ ID NO: 1-3 or referenced in Table A.

48. The composition of claim 47, wherein the RNA molecule is a short inhibitory RNA molecule.

Abstract of the Disclosure

Provided herein are methods for identifying a risk of osteoarthritis in a subject, reagents and kits for carrying out the methods, methods for identifying candidate therapeutics for treating osteoarthritis, and therapeutic and preventative methods applicable to osteoarthritis. These embodiments are based upon an analysis of polymorphic variations in nucleotide sequences within the human genome.

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

FIGURE 1A

KIAA0296 – DISCOVERY P-VALUES (female only)

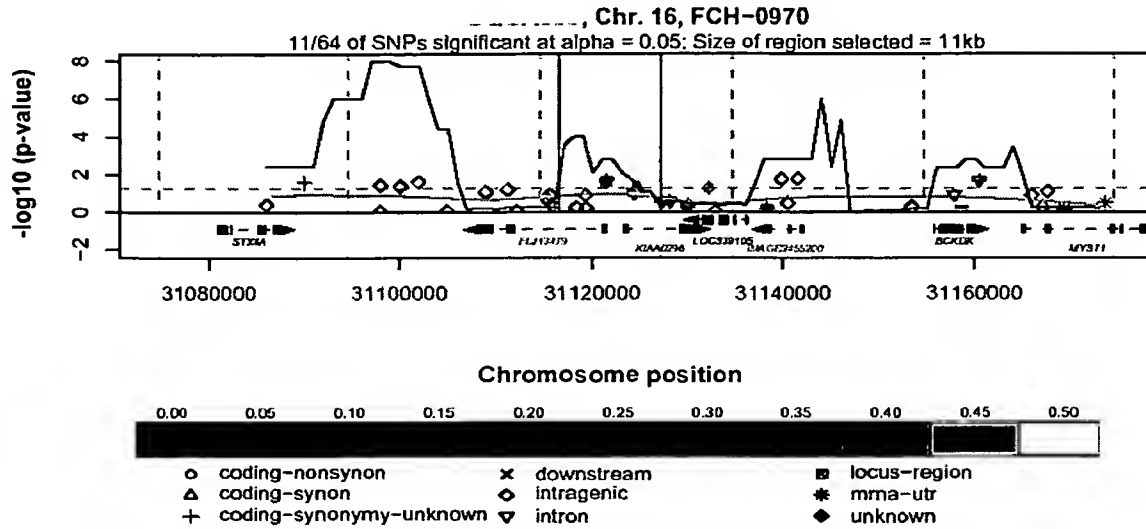


FIGURE 1B

CHROM 4 – DISCOVERY P-VALUES (female only)

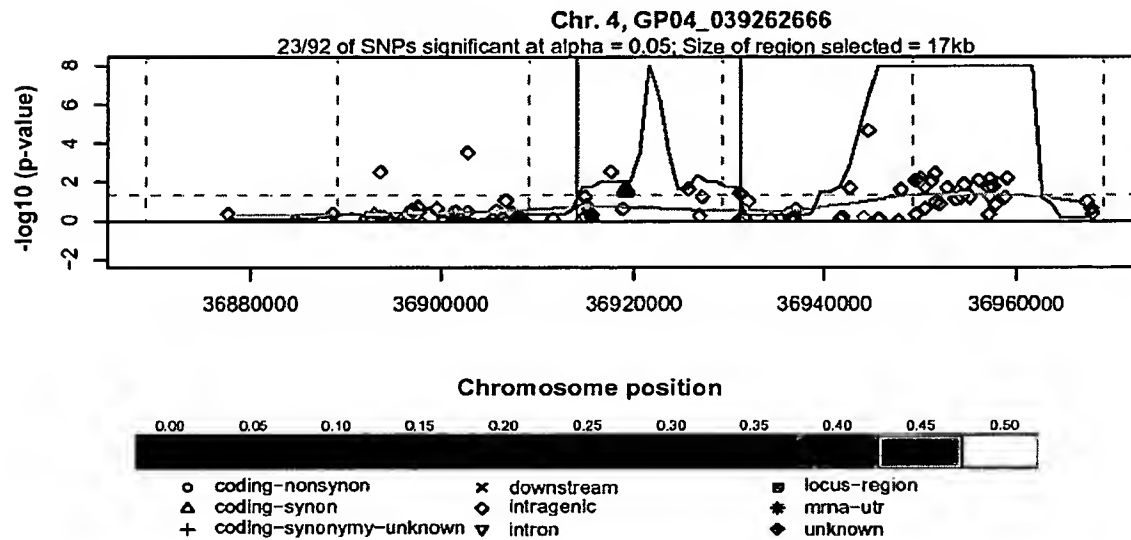
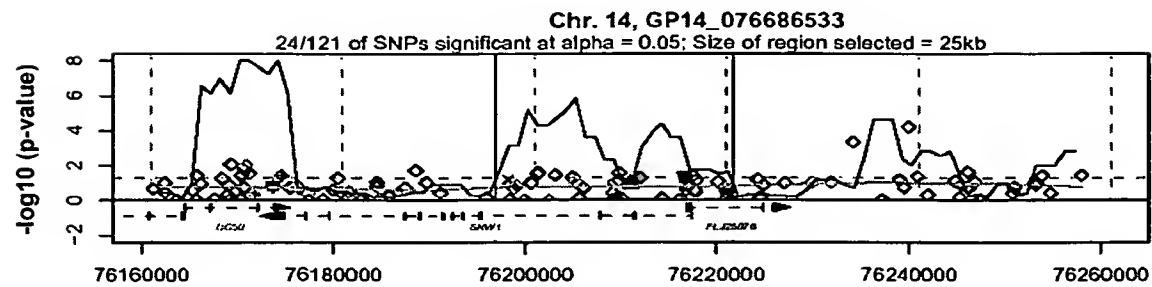
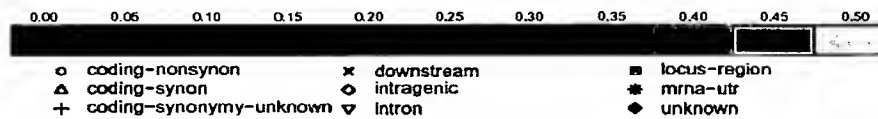


FIGURE 1C

SNW1 – DISCOVERY P-VALUES (female only)



Chromosome position



Application Data Sheet

Application Information

Application Type::	Provisional
Subject Matter::	Utility
Suggested Group Art Unit::	Not Yet Assigned
CD-ROM or CD-R?::	None
Sequence submission?::	None
Computer Readable Form (CRF)?::	No
Title::	METHODS FOR IDENTIFYING RISK OF OSTEOARTHRITIS AND TREATMENTS THEREOF
Attorney Docket Number::	524593008600
Request for Early Publication?::	No
Request for Non-Publication?::	No
Total Drawing Sheets::	2
Small Entity?::	Yes
Petition included?::	No
Secrecy Order in Parent Appl.?::	No

Applicant Information

Applicant Authority Type::	Inventor
Primary Citizenship Country::	US
Status::	Full Capacity
Given Name::	Steven
Family Name::	MAH
City of Residence::	San Diego
State or Province of Residence::	CA
Country of Residence::	US
Street of mailing address::	12820 Via Nieve #74
City of mailing address::	San Diego
State or Province of mailing address::	CA

Postal or Zip Code of mailing address:: 92130

Applicant Authority Type:: Inventor
Primary Citizenship Country:: Germany
Status:: Full Capacity
Given Name:: Andreas
Family Name:: BRAUN
City of Residence:: San Diego
State or Province of Residence:: CA
Country of Residence:: US
Street of mailing address:: 3935 Lago Di Grata Circle
City of mailing address:: San Diego
State or Province of mailing address:: CA
Postal or Zip Code of mailing address:: 92130

Applicant Authority Type:: Inventor
Primary Citizenship Country:: Germany
Status:: Full Capacity
Given Name:: Stefan
Middle Name:: M.
Family Name:: KAMMERER
City of Residence:: San Diego
State or Province of Residence:: CA
Country of Residence:: US
Street of mailing address:: 3825 Elijah Court, Unit 334
City of mailing address:: San Diego
State or Province of mailing address:: CA
Postal or Zip Code of mailing address:: 92130

Applicant Authority Type:: Inventor
Primary Citizenship Country:: US
Status:: Full Capacity

Given Name:: Matthew
Middle Name:: Roberts
Family Name:: NELSON
City of Residence:: San Marcos
State or Province of Residence:: CA
Country of Residence:: US
Street of mailing address:: 1250 Calle Prospero
City of mailing address:: San Marcos
State or Province of mailing address:: CA
Postal or Zip Code of mailing address:: 92069

Applicant Authority Type:: Inventor
Primary Citizenship Country:: Sweden
Status:: Full Capacity
Given Name:: Rikard
Middle Name:: Henry
Family Name:: RENELAND
City of Residence:: San Diego
State or Province of Residence:: CA
Country of Residence:: US
Street of mailing address:: 7555 Charmant Drive, #1114
City of mailing address:: San Diego
State or Province of mailing address:: CA
Postal or Zip Code of mailing address:: 92122

Applicant Authority Type:: Inventor
Primary Citizenship Country:: United Kingdom
Status:: Full Capacity
Given Name:: Maria
Middle Name:: L.
Family Name:: LANGDOWN
City of Residence:: San Diego

State or Province of Residence:: CA
Country of Residence:: US
Street of mailing address:: 3701 Yosemite Street
City of mailing address:: San Diego
State or Province of mailing address:: CA
Postal or Zip Code of mailing address:: 92109

Correspondence Information

Correspondence Customer Number:: 25225

Representative Information

Representative Customer Number:: 25225